NASA CONTRACTOR REPORT

NASA CR-193

RESEARCH IN AND APPLICATION OF MODERN AUTOMATIC CONTROL THEORY TO NUCLEAR ROCKET DYNAMICS AND CONTROL

by Lynn E. Weaver, Donald G. Schultz, David L. Hetrick, James L. Melsa, Hugh S. Murray, and Phillip A. Secker

Prepared under Grant No. NsG-490 *by*
UNIVERSITY OF ARIZONA
Tucson, Ariz.
for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MARCH 1965

RESEARCH IN AND APPLICATION OF MODERN AUTOMATIC CONTROL THEORY TO NUCLEAR ROCKET DYNAMICS AND CONTROL

By Lynn E. Weaver, Donald G. Schultz, David L. Hetrick,
James L. Melsa, Hugh S. Murray, and Phillip A. Secker

Prepared under Grant No. NsG-490 by
UNIVERSITY OF ARIZONA
Tucson, Ariz.

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Office of Technical Services, Department of Commerce,
Washington, D.C. 20230 -- Price $5.00

TABLE OF CONTENTS

INTRODUCTION

This progress report discusses the research performed under NASA Grant NsG-490 covering the period from February 1 to September 1, 1964. The purpose of this grant is the application and extension of modern automatic control theory to nuclear rocket dynamics and control. The report is composed of four independent sections each covering a specific part of the research program.

Section I "Closed-Loop Sub-Optimal Control Employing the Second Method of Liapunov" presents a new approach to the synthesis problem. It is an attempt to combine the Second Method of Liapunov and Pontryagin's Maximum Principle and results in a closed-loop control, as compared to open-loop control obtained using the Maximum Principle. The research to date has been concerned with developing the foundation for this approach. The work to be performed during the next report period will be devoted to the applications of this concept to the control of bilinear nuclear rocket system.

Section II "Nonlinear Stability of Coupled Core Reactor" is a study of the application of the Second Method of Liapunov to the stability of clustered nuclear rocket engines. Previous stability analysis of coupled core systems has been based on linear reactor models, with approximations to the neutron transport delay times. The purpose of this research is to consider more realistic system models with true delay times and to determine regions of stability.

Section III "Synthesis of Optimal Closed-Loop Control for Nuclear Rocket Systems" considers the realization of a closed control for an optimal control law determined by the Maximum Principle. In physical systems various

types of disturbance are encountered which makes open-loop control impractical. This section treats the determination of a closed loop control in the presence of measurement noise and external disturbances. Several approaches to the problem are discussed and an example worked.

Section IV "Limits of Validity for Some Approximations in Reactor Dynamics" evolved as a side interest in the program. It was felt to be of sufficient importance to be included in the report. This section discusses various approximations to the response of a reactor to a constant rate of reactivity increase. Such approximations are of particular interest during start-up.

Another part of the research program for which insufficient progress has been made to warrant a report is the work on the stability of loosely-coupled higher order system. This phase of the program recently begun is concerned with the use of the Second Method in determining regions of stability of such systems. It will be some time before definite information can be obtained.

SECTION I

CLOSED-LOOP SUB-OPTIMAL CONTROL EMPLOYING THE SECOND METHOD OF LIAPUNOV

Chapter 1

INTRODUCTION AND ORGANIZATION

1.1 Introduction

The problem of controlling a system such that its performance approximates in some sense a desired performance has been important for a long time. A natural outgrowth of this interest is the optimal control problem: controlling a system in such a manner that its performance is the best possible.

Within the last few years, several, rather elegant, general methods of solving the optimal control problem have been presented. Notable among these is the maximum principle of Pontryagin. In general, these methods involve unwieldly computations for all but trivial problems. Also in many cases, the control once obtained is of an open-loop nature, that is, valid for only one initial condition and no disturbances.

The difficulties associated with these methods have led to a growing gap between theoretical and practical control work. To fill this gap, there has been an ever-increasing development of special techniques for special problems which generally lead to sub-optimal control, control which is acceptably close to the true optimal but practicable.

In this work, the Second Method of Liapunov is used as a basis for developing such a method for closed-loop optimal control of linear systems with a bounded control norm. This method centers on the solution of a partial differential equation which is equivalent to the Hamilton-Jacobi equation. A special class of solutions, called eigenvector scalar

products, is shown to exist. These solutions are combined to form a sub-optimal control method which provides a practical compromise between system complexity and speed of response.

However, the development of this method is, at present, still incomplete. Therefore the material in this report is basically of a background nature and hence limited in its usefulness. Current research is pointed toward extending the approach in order to make it apply to a wider range of problems and hence to increase its usefulness. Preliminary results have indicated that this attempt should be very successful. Several future research topics are discussed in the last chapter.

1.2 Organization of the Report

This report consists of three basic parts. The first part comprising Chapter 1, 2, and 3 is introductory in nature. Following the introductory material in this chapter, the basic optimization problem to be considered is formulated in Chapter 2. Chapter 2 also contains a brief review of a modified form of the maximum principle which has been termed the minimum principle. In Chapter 3 a brief introduction to the Second Method of Liapunov is presented in order to make the work a self-contained unit.

Chapters 4 and 5 form the second part, the theoretical heart. In Chapter 4, the Second Method is combined with the minimum principle to develop another approach to the basic optimization problem. It is demonstrated that solving the basic optimization problem is equivalent to solving a first-order partial differential equation which is identical to the Hamilton-Jacobi equation. Although no general method of solving

this equation is known, a special class of solutions is shown to exist. This class of solutions, called eigenvector scalar products, is developed and discussed in detail in Chapter 5.

The third part, consisting of Chapter 6, is the practical portion of the work. In Chapter 6, the eigenvector scalar product solutions are combined to form an effective sub-optimal control method for systems in which the control matrix is non-singular. In this form, the sub-optimal control method provides an effective solution to a limited class of practical systems.

Chapter 7 contains a discussion of the concepts introduced and several ideas for further research. Examples are presented throughout the work whenever they can serve to better illustrate a point.

A basic knowledge of vector and matrix algebra is expected of the reader, as well as an understanding of the state variable method of formulating control problems. Although a brief review of the minimum principle and the Second Method are presented, the reader who is not familiar with these methods may wish to consult some of the suggested references for a more introductory presentation.

Chapter 2

MINIMUM PRINCIPLE

2.1 Introduction

This chapter consists of two basic parts. First, the basic optimization problem of this work is formulated, including all necessary definitions and notation. Second, a brief description of the minimum principle method for solving this problem is presented. Since extensive accounts of this method may be found in the literature[3,4,5], only the aspects pertinent to the particular problem of this work are included. Those familiar with the minimum principle may wish to skip section 2.4.

The chapter concludes with a short discussion of the inadequacy of the minimum principle approach in solving the optimization problem.

2.2 Notation

In this section, the notation which will be used throughout is explained. In general, the state space approach will be employed, utilizing vector-matrix formulation. Vectors will be indicated by lower case Roman letters such as x, u. One exception to this rule will be the letter t, which will indicate time, a scalar. The components of a vector will be indicated by subscripted lower case Roman letters, therefore $x = (x_1, x_2, \ldots, x_n)$. Particular vectors will be indicated by superscripts, therefore $x^1 = (x_1^1, x_2^1, \ldots, x_n^1)$.

Matrices will be designated by underlined upper case Roman letters such as $\underline{A}$, $\underline{B}$; scalars, by upper case Roman letters or Greek letters. The transpose of a matrix or vector will be designated by a prime, therefore x' is the transpose of the vector x.

The notation $\partial L(x)/\partial x$ will be used to indicate a vector whose components consist of the partial derivatives of $L(x)$, thus $\partial L(x)/\partial x = (\partial L(x)/\partial x_1, \ldots, \partial L(x)/\partial x_n)$. The notation $\nabla L(x)$ will also be used when the differentiation is with respect to x; thus, $\nabla L(x) = \partial L(x)/\partial x$.

2.3 Formulation of the Basic Optimization Problem

It will be assumed that the state of the control system can be completely described at any instant of time by n real numbers, $x_1, x_2, \ldots, x_n$. The behavior (or motion) of the system as a function of time may then be described by n real functions of time, $x_1(t), x_2(t), \ldots, x_n(t)$. These variables, called state variables, are the components of the state vector $x(t) = (x_1(t), x_2(t), \ldots, x_3(t))$.

It will be further assumed that the motion of the system can be controlled by a set of r real valued control variables, $u_1(t), u_2(t), \ldots u_r(t)$, which are the components of the control vector, $u(t)$. The set of all possible values of u is called the control region, U, a subset of a r-dimensional Euclidean space. In most practical applications, U is closed and bounded.

For the present work U will consist of the set of all u such that $\|\underline{D}u\|^2 \leq \alpha^2$ where $\underline{D}$ is a non-singular matrix and α is a real constant. However, by a simple change of variables $w = \alpha^{-1}\underline{D}u$, $\|\underline{D}u\|^2 \leq \alpha^2$ becomes $\|w\|^2 \leq 1$. Hence there is no loss of generality in considering $\underline{D}$ to be

the identity matrix and α to be unity. Thus U will be the set of all u such that $\|u\|^2 \leq 1$. If $u(t) \in U$ and is, in addition, piecewise continuous, then u(t) is called an admissible control.

The only systems to be considered here are ones for which the laws of motion may be written as a set of n first-order linear equations.

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij}x_j + \sum_{k=1}^{r} b_{ik}u_k \qquad i = 1,2,\ldots,n \tag{2.1}$$

Or written in vector-matrix notation

$$\dot{x} = \underline{A}x + \underline{B}u \tag{2.2}$$

It will be assumed that corresponding to every admissible control u(t) and every initial condition $x^o = x(t_o)$, that the motion of the system is defined uniquely by the solution of equation (2.2). This solution is called the solution (or motion) of the system corresponding to the control u(t) for the initial condition x^o.

An admissible control is said to transfer the system from x^o to x^1 if the solution corresponding to that control and the initial condition x^o is defined for $t_o \leq t \leq t_1$ and reaches x^1 at the time t_1.

Since, in general, there may be many admissible controls which transfer the system from x^o to x^1, the question which naturally arises is, "Which admissible control, in addition to transferring the system from x^o to x^1, minimizes some cost functional

$$J = \int_{t_o}^{t_1} L(x(t))dt \tag{2.3}$$

where L(x) is a real and positive-valued function of the state vector?"

It should be noted that for fixed points the transition time, $t_1 - t_o$, is not fixed but is dependent on the particular control used. One example of particular importance is the case when $L(x) = 1$ and the cost functional, J, reduces to $t_1 - t_o$, the transition time. This is the familiar time-optimal problem which is treated in detail in later chapters.

A control which transfers the system from x^o to x^1 while minimizing the cost functional is called an optimal control corresponding to a transition from x^o to x^1. For convenience, x^1 is considered to be the origin for the rest of this work.

The optimal control may be found in two different forms. First, the control variables may be obtained as functions of time during the transition interval $t_1 - t_o$ for a given initial condition x^o. This is called open-loop control, since no information concerning the system state is needed or used during the transition interval.

Second, the control variables may be determined as explicit functions of the system state, i.e., $u = u(x)$. This is called closed-loop control, since knowledge of the system state is used during the transition interval. The advantages of closed-loop control are well established in the literature[1,2] and therefore only three points are mentioned here. First, feedback or closed-loop operation reduces the effect of system parameter variations. Second, feedback operation minimizes the effect of external disturbances. Third, in many practical cases the equations of motion are known only approximately. By the use of closed-loop control, variations in the systems motion due to these inaccuracies can be minimized. Thus it appears obvious that not only should one seek optimal control, but, in general, one should seek closed-loop optimal control.

The fundamental problem may then be stated in the following form. Given a linear system whose laws of motion are described by equation (2.2), it is desired to find an optimal, closed-loop, admissible control corresponding to a transition from x^o to the origin with a cost functional of the form of equation (2.3). Additional assumptions concerning the system and the cost functional will be made in later chapters.

The next section presents the basic formulation and theorems of the minimum principle, a method for obtaining an open-loop solution of the above problem.

2.4 Minimum Principle

The concept of the minimum principle was first introduced by Kalman[3] as a minor modification of the maximum principle developed by Pontryagin and his students[4]. The essential differences between the two approaches are noted below. The minimum principle is a logical extension of the classical calculus of variations and provides a broad and unifying approach to a wide variety of variational and optimal control problems. Only those aspects of the theory which are pertinent to the problem of the preceding section are presented here.

As the first step in the minimum principle approach, a new set of n variables, p_i, are adjoined to the state variables, x_i, of the system. These new variables, called adjoint variables, are defined by the following set of differential equations, the adjoint equation.

$$\dot{p}_i = -\frac{\partial}{\partial x_i}\left(\sum_{m=1}^{n} p_m \dot{x}_m + L(x)\right) \qquad i = 1,2,\ldots,n \qquad (2.4)$$

Next a scalar function H analogous to the Hamiltonian is defined by

$$H(x,p,u) = p'\dot{x} + L(x) \tag{2.5}$$

It can be readily verified that equations (2.2) and (2.4) can be rewritten in terms of H(x,p,u) in the following system of equations which are analogous to the Hamiltonian canonic equations.

$$\begin{aligned} \dot{x}_i &= \frac{\partial}{\partial p_i}(H(x,p,u)) \\ \dot{p}_i &= \frac{-\partial}{\partial x_i}(H(x,p,u)) \quad i = 1,2,\ldots,n \end{aligned} \tag{2.6}$$

For fixed values of x and p, H becomes a function of the control vector u. The greatest lower bound of this function with respect to admissible controls $u \in U$ will be denoted by H^o, therefore

$$H^o(x,p) = \inf_{u \in U} H(x,p,u) \tag{2.7}$$

If the continuous function H actually assumes its lower bound on U, then H^o will be the minimum of H on U. This will be true for all problems in this work, hence

$$H^o(x,p) = \min_{u \in U} H(x,p,u) \tag{2.8}$$

The corresponding minimizing control will be designated by u^o.

The following theorem presents a necessary condition for the optimality of a control u.

Theorem 2.1 Let u(t), $t_o \leq t \leq t_1$, be an admissible control such that the corresponding motion x(t) which begins at the point x^o at time t_o reaches, at time t_1, the point x^1. In order that u(t)

and x(t) be optimal, it is necessary that there exist a nonzero continuous vector function p(t) corresponding to u(t) and x(t) such that:

1) for every t, $t_o \leq t \leq t_1$, the function $H^o(x,p,u)$ of the variable $u \in U$ attains its minimum at the point $u = u(t)$:

$$H(x,p,u) = H^o(x,p)$$

2) for every t, $t_o \leq t \leq t_1$, the function $H^o(x,p)$ is identically zero:

$$H^o(x(t),p(t)) = 0$$

This theorem formulated in terms of the minimum principle is equivalent to a theorem of the maximum principle initially proven by Pontryagin[5]. In the maximum principle formulation, the sign preceding L(x) in both equation (2.4) and (2.5) is negative. Because of this change, it is necessary to consider the least upper bound of H(x,p,u), rather than the greatest lower bound. Hence H is maximized rather than minimized. Although the use of the maximum principle is more common in the literature, the use of the minimum principle is more convenient for the development of Chapter 4 and thus it is employed here.

For the problem presented in the preceding section, the Hamiltonian is given by

$$\begin{aligned} H(x,p,u) &= p'(\underline{A}x + \underline{B}u) + L(x) \\ &= p' \underline{A}x + p' \underline{B}u + L(x) \end{aligned} \tag{2.9}$$

The adjoint equations (2.4) may then be developed by use of equation (2.6)

$$\dot{p} = -\underline{A}'p - \nabla L(x) \tag{2.10}$$

The next step is the minimization of $H(p,x,u)$ with respect to $u \in U$. Since the middle term on the right side of equation (2.9) is the scalar product of two vectors, $p'\underline{B}$ and u, $H(x,p,u)$ is minimized by making the direction of u opposite to $\underline{B}'p$ and making the magnitude of u as large as possible. However, the norm of u is required to be less than or equal unity in order for u to be an admissible control. Hence, u is selected to be a vector with unit norm (length) and direction opposite $\underline{B}'p$:

$$u^o = - \frac{\underline{B}'p}{\|\underline{B}'p\|} \tag{2.11}$$

Substituting u as given by equation (2.11) into the equations (2.2) and (2.9), the following set of coupled first-order ordinary differential equations are obtained.

$$\dot{x} = \underline{A}x + \frac{\underline{B}\underline{B}'p}{\|\underline{B}'p\|} \quad . \tag{2.12}$$

$$\dot{p} = -\underline{A}'p - \nabla L(x) \tag{2.13}$$

with the boundary conditions $x(t_o) = x^o$ and $x(t_1) = x^1$ and the auxiliary condition $H^o(x,p) = 0$.

The difficulties inherent in the minimum principle approach are now obvious. First, the simultaneous solution of equations (2.12) and (2.13) is not elementary, since both equations are in general nonlinear. The adjoint equations have no boundary conditions while the system equations have second boundary conditions which creates the so called "two-point" boundary value problem. Normally numerical solution of these equations is necessary. Second, the control as determined by the minimum principle is open-loop control, i.e., $u = u(t)$ not $u(x)$.

Another method for attacking the basic optimization problem of the preceding section is presented in Chapter 4. The method is based on both the Second Method of Liapunov and the minimum principle and attempts to remove or alleviate the difficulties mentioned above. In particular, the control vector is found as a function of the state variables, i.e., closed-loop control. However, before proceeding to that development, it is necessary to present some of the basic definitions and theorems of the Second Method.

Chapter 3

SECOND METHOD OF LIAPUNOV

3.1 Introduction

The Second Method of Liapunov provides the most general approach to the stability of dynamic systems whose laws of motion are described by ordinary linear or nonlinear differential equations. This chapter presents a brief review of the basic concepts and definitions of the Second Method. Only those portions of the theory which are directly applicable to the problem at hand will be discussed. The reader is directed to the literature for a more complete presentation[6,7,8,9].

In this chapter, the dynamic systems under consideration are assumed to be autonomous and describable in state variable form as n first-order differential equations of the form

$$\dot{x}_i = f_i(x) \qquad i = 1, 2, \ldots, n \tag{3.1}$$

In matrix notation, this may be written as

$$\dot{x} = f(x) \tag{3.2}$$

Such a system is called autonomous. It is obvious that for closed-loop control the system of equation (2.2) is of this form since it becomes

$$\begin{aligned} \dot{x} &= \underline{A}x + \underline{B}u(x) \\ &= f(x) \end{aligned}$$

The equilibrium state being investigated is assumed to be located at the origin. This is actually no restriction, since any equilibrium point may always be translated by simple linear change of variables to the origin. Again the system discussed in Chapter 2 satisfies this assumption since the control is always chosen such as to drive the system to the origin.

This chapter consists of three parts. First, the definitions of definiteness and stability are presented. Second, a modified Liapunov stability theorem is stated without proof. Third, this stability theorem is given a geometric interpretation.

3.2 Definitions

The concepts of definiteness play an important role in the stability theorems. The following definitions, which follow Malkin, are of interest here.

Definition 3.1 **Positive (Negative) Definite**

A scalar function, $V(x)$, is positive (negative) definite if for $\|x\| \leq \alpha$ $V(x) > 0$ (<0) for all $x \neq 0$ and $V(0) = 0$.

Definition 3.2 **Positive (Negative) Semidefinite**

A scalar function, $V(x)$, is positive (negative) semidefinite if for $\|x\| \leq \alpha$ $V(x) \geq 0$ (≤ 0) for all $x \neq 0$ and $V(0) = 0$.

Definition 3.3 **Indefinite**

A scalar function, $V(x)$, is indefinite if no matter how small α is chosen, $V(x)$ may assume both positive and negative values for $\|x\| \leq \alpha$.

If in the above definitions α may be made arbitrarily large, in which case the definitions hold in the whole space. This will be the case with all of the scalar functions to be discussed in the following chapters.

A few examples will serve to clarify these definitions. The function

$$V(x) = (x_1)^2 + (x_2)^4$$

is positive definite if the system is second-order, but is only semi-definite if the system is of higher order, since for $x_1 = x_2 = 0$, $V(x)$ will be zero independent of x_3, x_4,... . On the other hand the function

$$V(x) = (x_1 + x_2)^2$$

is semidefinite even for second-order systems, since if $x_1 = -x_2$, $V(x)$ will be zero even though x is not equal to zero. The function

$$V(x) = x_1 + x_2$$

is obviously indefinite independent of the order of the system.

One class of scalar functions that will be particularly important is a quadratic form. In this case $V(x)$ may be written in the form

$$V(x) = x'\underline{C}x$$

where $\underline{C}$ is a constant square matrix. Usually if $V(x)$ is a quadratic form, the definiteness of $V(x)$ is attributed to $\underline{C}$. Hence one speaks of a positive definite matrix.

Closely related to the concept of definiteness is the concept of a simple closed surface (or curve). A surface is said to be simple if it does not intersect itself and closed if it intersects all paths that lead from the origin to infinity. The reader is reminded that it is assumed that the equilibrium state is at the origin. Hence a simple closed surface is topologically equivalent to the surface of an n-dimensional sphere. Letov[11] has shown that if a scalar function, $V(x)$,

is positive definite and, in addition, is radially unbounded, i.e., $V(x) \rightarrow \infty$ as $\|x\| \rightarrow \infty$, then the set of all points x such that $V(x) = K$, a positive constant, is a simple closed surface. In addition, the surface $V(x) = K_1$ lies entirely inside the surface $V(x) = K_2$ whenever $K_1 > K_2$.

There are many types of stability that have been defined for systems that may be described by equation (3.2). In the case of linear systems, almost all of these definitions are equivalent. For nonlinear systems, this is not true. However for this work, only stability in the sense of Liapunov and asymptotic stability are of interest. Hence only these types of stability are defined. Let $S(\alpha)$ be the spherical region of radius $\alpha > 0$ around the origin, i.e., $S(\alpha)$ consists of all points x such that $\|x\| < \alpha$.

Definition 3.4 Stable in the Sense of Liapunov

The origin is stable in the sense of Liapunov, or simply stable, if corresponding to every number $\epsilon > 0$ there exists a number $\delta(\epsilon) > 0$ such that solutions starting in $S(\delta)$ will remain in $S(\epsilon)$ ever after.

Definition 3.5 Asymptotically Stable

If the origin is stable and, in addition, every solution starting in $S(\delta)$ not only stays in $S(\epsilon)$ but tends toward the origin as time increases indefinitely, then the origin is asymptotically stable.

Definition 3.6 Unstable

The origin is unstable if for some $\epsilon > 0$ and any $\delta > 0$, no matter how small, there is always a point x in $S(\delta)$ such that a solution starting from that point leaves $S(\epsilon)$.

A graphical representation of these definitions is shown in Figure 3.1 for a two-dimensional case.

The definitions emphasize the local character of stability for nonlinear systems, since the region $S(\delta)$ may be arbitrarily small. If

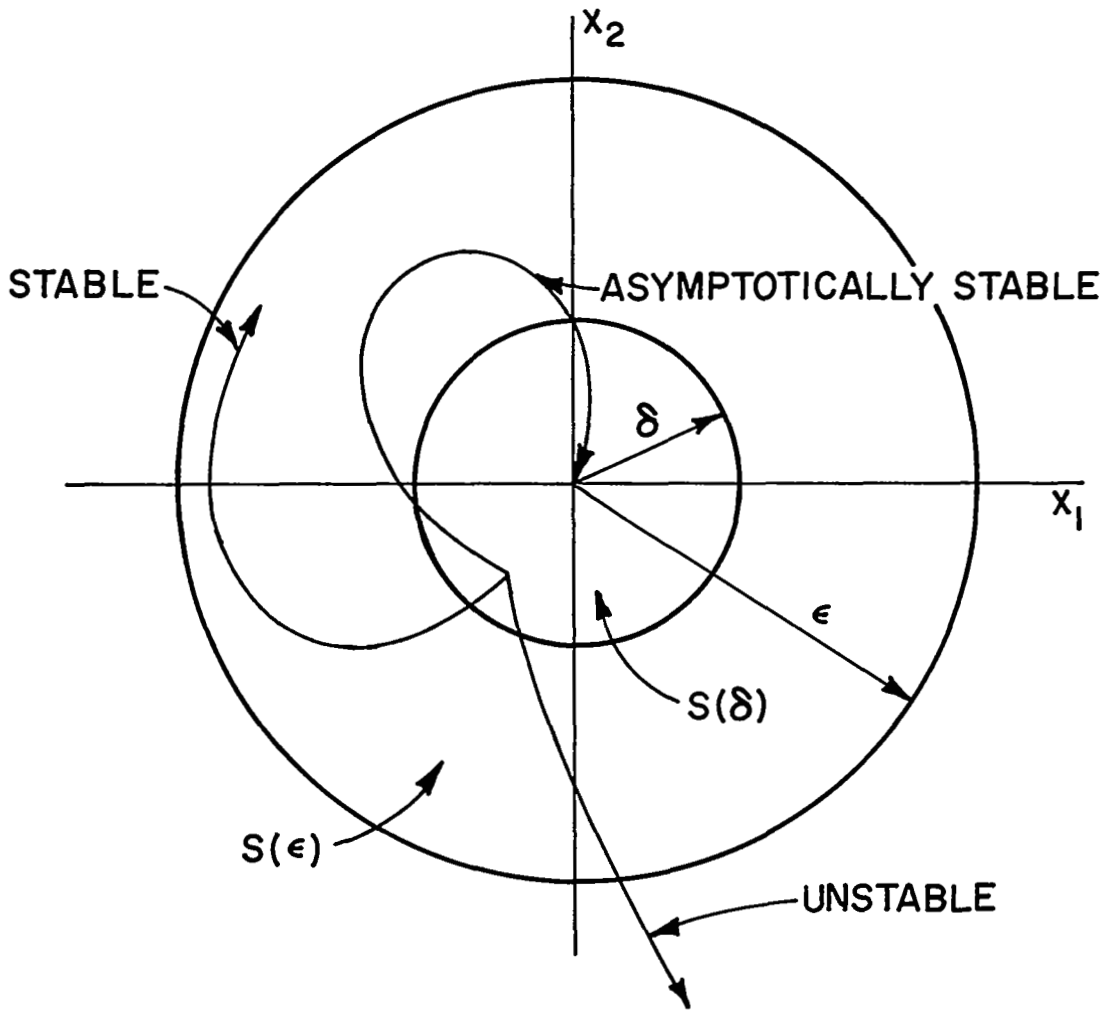


FIGURE 3.1 GRAPHICAL REPRESENTATION OF STABILITY DEFINITIONS

the region $S(\delta)$ includes the entire space, the definitions are called global. In the chapters which follow the main interest is in global asymptotic stability, since the systems are linear.

3.3 Stability Theorem

As was the case with definitions of stability, there are many stability theorems which constitute the Second Method. Since the major concern of this work is not stability, only one theorem is presented here. This stability theorem, due to LaSalle[7], differs from the original Liapunov theorem in the first condition where $\dot{V}(x)$ is allowed to be semidefinite, as long as it is not zero on a solution of the system, other than the origin. In the original theorem, $\dot{V}(x)$ was required to be negative definite.

Theorem 3.1 Stability Theorem If there exists a positive definite scalar function $V(x)$ with continuous first partials such that

1) $\dot{V}(x) \leq 0$ for all x (at least negative semidefinite)

2) $V(x) \rightarrow \infty$ as $\|x\| \rightarrow \infty$ (radially unbounded)

then if $\dot{V}(x)$ is not identically zero along any solution of (3.2) other than the origin, the system is globally asymptotically stable.

Since $V(x)$ has continuous first partials, the chain rule may be used to obtain $\dot{V}(x)$

$$\dot{V}(x) = \frac{dV(x)}{dt} = \frac{\partial V(x)}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial V(x)}{\partial x_2}\frac{dx_2}{dt} + \ldots + \frac{\partial V(x)}{\partial x_n}\frac{dx_n}{dt}$$

$$= \sum_{i=1}^{n} \frac{\partial V(x)}{\partial x_i}\dot{x}_i$$

which may be written with the use of the notation $\nabla V(x)$ as

$$\dot{V}(x) = \nabla V'(x)\dot{x} \tag{3.3}$$

The basic concept of the Second Method is now evident: by proper selection or generation of a Liapunov V-function, it is possible to determine the stability of a nonlinear dynamic system without any knowledge of the solutions of the system equation. It is perhaps of value to investigate the stability theorem from a geometric viewpoint.

Since $V(x)$ is positive definite, and radially unbounded $V(x) = K$, a constant, becomes a family of concentric closed surfaces surrounding the origin such that the surface $V(x) = K_1$ lies inside $V(x) = K_2$ whenever $K_1 > K_2$. Figure 3.2 shows a graphical picture for the two-dimensional or second-order case. Since both $V(x)$ and $\dot{V}(x)$ are implicit functions of time and $\dot{V}(x)$ is required to be non-positive, the state of the system must be found on successively "smaller" $V(x) = K$, a constant, surfaces or must remain stationary. But $\dot{V}(x)$ cannot be zero on any solution except $x = 0$; therefore the state of the system cannot remain stationary. Hence, the system trajectory must move toward the origin.

Three features of the Second Method should be noted. First, the method provides only sufficient conditions for stability; hence if a system does not satisfy the stability theorem, no conclusion may be drawn relative to system stability. Second, the converse of the stability theorem has been proven. Therefore if the system is stable, a V-function must exist. Third, the V-function is not unique, which is one of the most powerful features of the Second Method. No longer is one searching for a single unique solution to the differential equation but rather for one,

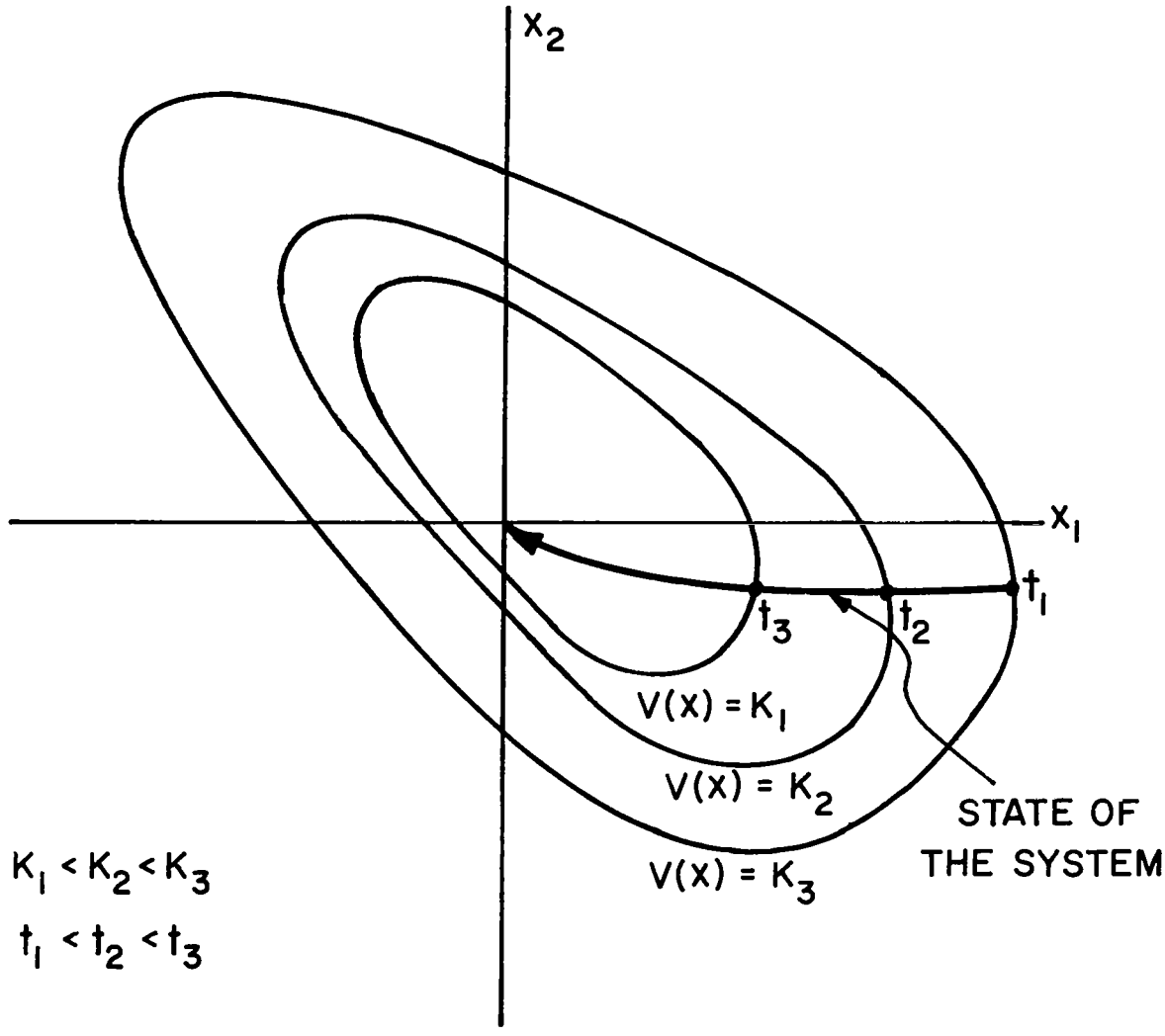


FIGURE 3.2 SURFACES OF V(x) = CONSTANT

out of many, V-functions. However because the method provides only sufficient conditions, some V-functions may provide a better answer than others.

Chapter 4

CLOSED-LOOP OPTIMAL CONTROL VIA THE SECOND METHOD

4.1 Introduction

In Chapter 2 the basic optimization problem was presented. This was followed by one method of obtaining an open-loop solution of the problem, the minimum principle. In this chapter another method of attacking the basic optimization problem is presented. This method, based on the Second Method of Liapunov and the minimum principle, yields closed-loop control.

In the next section a brief discussion of the background for the use of the Second Method is presented. This is followed by two optimality theorems and their proofs. It is demonstrated that solving the basic optimzation problem is equivalent to solving a first-order partial differential equation which is identical to the Hamilton-Jacobi equation. Since no general method of solving this equation is known, the approach presented here has not solved the problem but has rather formulated the problem into a new framework. In this framework, a special class of solutions, called eigenvector scalar products, is shown to exist in the next chapter. From these solutions, a method for designing effective closed-loop, sub-optimal control is developed.

It should be noted that the results of this chapter are not new, although the method of deriving them is. As is shown in the last section of this chapter, the results could have been derived directly from the

Hamilton-Jacobi equation. In effect, a special case of the Hamilton-Jacobi equation is derived in this chapter. It was felt that carrying out the development in this manner adds greater insight into the relation between the Second Method and optimal control.

4.2 Background

The use of the Second Method of Liapunov for the design of optimal systems has been suggested by several authors[11,12,13,14,15]. Unfortunately, almost all of these methods have three basic problems: 1) they are approximate, 2) either no estimate of the approximation error is possible, or the estimate is overly conservative, and 3) it is necessary to choose a V(x) for which no general procedure is presented. Hence these methods were never widely accepted. (A brief resume of several of these methods can be found in the Appendix.)

Nahi[14] has recently presented a procedure for using the Second Method to obtain time-optimal control. However, Nahi was only able to find solutions for a rather restricted class of systems. It is shown later that Nahi's method is a special case of the method presented here.

The determination of $\dot{V}$ from V(x) was discussed in Chapter 3; the result is repeated here for reference.

$$\dot{V} = \nabla W'(x)\dot{x} \tag{4.1}$$

Now substituting equation (2.2) for $\dot{x}$, one obtains

$$\dot{V} = \nabla W'(x)\underline{A}x + \nabla W'(x)\underline{B}u \tag{4.2}$$

Thus $\dot{V}$ becomes a function of both the control and state vectors for a given V(x). In the following discussion the notation $\dot{V}(x,u)$ will be used

to indicate this dependence on both u and x.

In 1960 Kalman and Bertram[12] presented a method for designing approximately time-optimal control systems. Their method was based on the knowledge that for a closed, bounded control region, the control vector is always on the boundary. They suggested minimizing $\dot{V}(x,u)$ with respect to all admissible controls based on the argument that this would make V(x) approach zero most rapidly and hence the system would reach the origin in minimum time. This method suffers from all of the disadvantages noted above and therefore has not been widely employed. However, the concept of minimizing $\dot{V}(x,u)$ is valuable and is used below.

Retaining the idea of minimizing $\dot{V}(x,u)$ for the moment, consider the implication of setting $\dot{V}(x) = -L(x)$. Since L(x) was required to be at least positive semidefinite, $\dot{V}(x)$ will thus be of the proper nature. Then V(x) becomes equivalent to the cost functional:

$$V(x(t_1)) - V(x(t_o)) = \int_{t_o}^{t_1} \dot{V}(x)dt = \int_{t_o}^{t_1} -L(x)dt \qquad (4.3)$$

Hence surfaces of constant V(x) become surfaces of constant cost.

The combination of these two concepts suggests the idea of setting $\min_{u\in U} \dot{V}(x,u) = -L(x)$. The question remaining is "Does this provide optimal control?" The following section demonstrates that the answer is affirmative.

Before proceeding with the proof in the next section, it should be pointed out that all of the approaches employing the Second Method yield closed-loop control. This is a feature that cannot be over-emphasized.

4.3 Optimality Theorems

In the preceding section, it was suggested that the selection of a V-function, $V(x)$, such that $\min_{u \in U} \dot{V}(x,u) = -L(x)$, would yield optimal control. In this section, corresponding optimality theorems are stated and their proofs given.

Before doing this it is perhaps of value to state the basic optimization problem again. Given a linear system whose laws of motion can be described by

$$\dot{x} = \underline{A}x + \underline{B}u$$

it is desired to find an optimal, closed-loop, admissible control corresponding to a transition from x^o to the origin with a cost functional of the form

$$J = \int_{t_o}^{t_1} L(x(t))dt$$

The control region, U, is the set of all control vectors, u, such that $\|u\|^2 \leq 1$.

For fixed values of x, $\dot{V}(x,u)$ becomes a continuous function of u. The minimum of this function with respect to all admissible control is designated by $\dot{V}^o(x)$.

$$\dot{V}^o(x) = \min_{u \in U} \dot{V}(x,u) \tag{4.4}$$

Anticipating the results to follow, the corresponding minimizing control is again denoted by u^o.

Theorem 4.1 If there exists a Liapunov function, $V(x)$, with continuous second partial derivatives with respect to x and such that $\dot{V}^o(x) = -L(x)$, then the control u^o which minimizes $\dot{V}(x,u)$ is an optimal control.

Before carrying out the proof of this theorem, consider the following lemma.

Lemma 4.1 If there exists a Liapunov function, V(x), with continuous second partial derivatives with respect to x and such that $\dot{V}^{o}(x) = -L(x)$, then the gradient of V(x), $\nabla W(x)$, satisfies the adjoint equation (2.4).

The first step in the proof of the lemma is the minimization of $\dot{V}(x,u)$ as given by equation (4.2) with respect to all admissible controls. The only term involving u is a scalar product of u and $\underline{B}'\nabla W(x)$. Thus by an argument similar to that presented in section 2.4, u^{o} is found to be

$$u^{o} = \frac{-\underline{B}'\nabla W(x)}{\|\underline{B}'\nabla W(x)\|} \tag{4.5}$$

Substituting u^{o} for u in equation (4.2), one obtains

$$\dot{V}^{o} = \nabla W'(x)\underline{A}x - \|\underline{B}'\nabla W(x)\| \tag{4.5A}$$

Setting $\dot{V}^{o}(x) = -L(x)$ yields

$$\nabla W'(x)\underline{A}x - \|\underline{B}'\nabla W(x)\| = -L(x) \tag{4.6}$$

Now taking the partial derivative of both sides of equation (4.6) with respect to x gives

$$\nabla(\nabla W'(x))\underline{A}x + \underline{A}'\nabla W(x) - \frac{\nabla(\nabla W'(x))\underline{B}\underline{B}'\nabla W(x)}{\|\underline{B}'\nabla W(x)\|} = -\nabla L(x) \tag{4.7}$$

Therefore

$$\nabla(\nabla W'(x))\underline{A}x = -\underline{A}'\nabla W(x) + \frac{\nabla(\nabla W'(x))\underline{B}\underline{B}'\nabla W(x)}{\|\underline{B}'\nabla W(x)\|} - \nabla L(x) \tag{4.8}$$

But from equation (4.5)

$$u^o = \frac{-\underline{B}'\nabla V(x)}{\| \underline{B}'\nabla V(x)\|}$$

and hence equation (4.8) becomes

$$\nabla(\nabla V'(x))\underline{A}x = -\underline{A}'\nabla V(x) - \nabla(\nabla V'(x))\underline{B}u_o - \nabla L(x) \tag{4.9}$$

Now consider the total time derivative of $\nabla V(x)$, again using the chain rule:

$$\frac{d}{dt}(\nabla V(x)) = (\nabla(\nabla V'(x)))'\dot{x} = (\nabla(\nabla V'(x)))'\underline{A}x + (\nabla(\nabla V'(x)))'\underline{B}u_o \tag{4.10}$$

By hypothesis $V(x)$ has continuous second partial derivatives, and therefore the matrix $\nabla(\nabla V'(x))$ is symmetric. Thus $\nabla(\nabla V'(x)) = (\nabla(\nabla V'(x)))'$. Then substituting equation (4.9) into equation (4.10) one obtains

$$\frac{d}{dt}(\nabla V(x)) = -\underline{A}'\nabla V(x) - \nabla L(x) \tag{4.11}$$

Comparing equation (4.11) with the adjoint equation (2.13), one notes that $\nabla V(x)$ satisfies the adjoint equation, which completes the proof of the lemma.

Now returning to the proof of theorem 4.1, $\nabla V(x)$ is substituted for p in the Hamiltonian as defined by equation (2.9) to obtain

$$\begin{aligned} H(x,\nabla V(x),u) &= \nabla V'(x)\underline{A}x + \nabla V'(x)\underline{B}u + L(x) \\ &= \overset{\circ}{V}(x,u) + L(x) \end{aligned} \tag{4.12}$$

Since $L(x)$ is not an explicit function of u,

$$\min_{u\in U} H(x,\nabla V(x),u) = \min_{u\in U} \dot{V}(x,u) + L(x)$$

Or

$$H^o(x,\nabla V(x)) = \overset{\circ}{V}{}^o(x) + L(x) \tag{4.13}$$

But by hypothesis, $\dot{V}^o(x) = -L(x)$ and hence

$$H^o(x,\nabla V(x)) = 0 \tag{4.14}$$

Therefore conditions 1 and 2 of the minimum principle have been satisfied and u^o must be an optimal control, which completes the proof of theorem 4.1. Again it should be noted that the control given by equation (4.5) is a closed-loop control. This theorem is discussed further in the next section.

The following theorem indicates an additional relationship between the Second Method and optimal control.

Theorem 4.2 If there exists a Liapunov function V(x) with continuous second partial derivatives such that $\nabla V(x)$ satisfies the adjoint equation and if u^o is an optimal control, then u^o minimizes $\dot{V}(x,u)$ and $\dot{V}^o(x) = -L(x)$.

Since u^o is an optimal control, it must minimize $H(x,p,u)$. But $\nabla V(x)$ satisfies the adjoint equation and hence it can be substituted for p. Then u^o must also minimize $H(x,\nabla V(x),u)$. By reference to equation (4.12), it can be concluded that u^o must also minimize $\dot{V}(x,u)$ since L(x) is not a function of u^o.

An application of the second condition of the minimum principle gives

$$\min_{u \in U} H(x,\nabla V(x),u) = H^o(x,\nabla V(x)) = 0$$

But by use of equation (4.13)

$$\dot{V}^o(x) = H^o(x,\nabla V(x)) - L(x) = -L(x)$$

and the theorem is proven.

In the next section, theorems 4.1 and 4.2 are discussed further, in particular with respect to the classical Hamilton-Jacobi equation. The existence of Liapunov functions as required for these theorems is also discussed.

4.4 Hamilton-Jacobi Equation

It was demonstrated, in the previous section, that the optimal control problem with a constraint on the norm of the control vector is equivalent to the problem of solving the first-order partial differential equation

$$\dot{V}^{o}(x) = -L(x) \tag{4.15}$$

It is of interest to note that equation (4.15) is, in fact, a special case of the classical Hamilton-Jacobi equation. The Hamilton-Jacobi equation may be obtained by setting $H^{o}(x,\nabla W(x)) = 0$. Thus for the problem of section 2.3, one obtains

$$H^{o}(x,\nabla W) = \nabla W'(x)\underline{A}x - \| \underline{B}'\nabla W(x)\| + L(x) = 0$$

or

$$H^{o}(x,\nabla W(x)) = \dot{V}^{o}(x) + L(x) = 0$$

Use could have been made of this fact in the development of the previous section. However, it was felt that greater insight into the use of the Second Method was obtained by carrying out the proof in the manner presented. The knowledge that equation (4.15) is the Hamilton-Jacobi equation does make it possible to conclude that the existence of a solution of equation (4.15) is sufficient for optimal control to exist. This is an advantage over the minimum principle where only necessary conditions for optimality are given.

Next, one might ask if solutions of sufficient smoothness, i.e., continuous second partial, derivatives exist for equation (4.15). Since the solutions of interest in the following chapter do, ipso facto, exist, the existence of solutions is not of prime importance here. However, it is perhaps of interest to look briefly at the problem, even though a complete answer is not known.

First, it can be shown by example that if the control is scalar and the system is at least second-order, then there is no solution of sufficient smoothness. In fact, there is no solution with continuous first partial derivatives. On the other hand, Krassovskii[16] has shown that if $\underline{B}$ is non-singular and $L(x) = 1$, then a solution to equation (4.15) exists with continuous partial derivatives of all order.

Hence, one is faced with a two-fold problem. First, a solution may not exist; and second, if one does exist, no general method of obtaining it is known. Therefore the basic optimization problem has not been solved. The necessary course of action is to obtain an approximate solution. In the next chapter, a method for modifying the Hamilton-Jacobi equation is followed by the presentation of a special class of solutions. From these solutions, a method for designing effective sub-optimal control is developed.

Chapter 5

EIGENVECTOR SCALAR PRODUCT SOLUTIONS

5.1 Introduction

In this chapter, a special class of solutions of the Hamilton-Jacobi equation is shown to exist. These solutions, called eigenvector scalar products, comprise the first of the three major contributions of this work. The second major contribution, which is contained in the last section of this chapter, is the development of a method for obtaining surfaces which bound the optimal isochrones from the outside. The next chapter forms the third major contribution, a method of designing effective sub-optimal control systems by the use of the eigenvector scalar product solutions.

The first part of this chapter presents a method of modifying the Hamilton-Jacobi equation in order to put the solution into a more convenient form. This is followed by the presentation of the eigenvector scalar product solutions. The last section of this chapter discusses the problem of bounding the optimum cost functional.

5.2 Modification of Hamilton-Jacobi Equation

A method of modifying the Hamilton-Jacobi equation is presented in this section which provides a more convenient representation of the solutions to be discussed in the next section. One approach might be to make a nonlinear transformation of coordinates in order to reduce the

Hamilton-Jacobi equation to some elementary form. To date this approach has not been very useful.

Another approach is to change to another Liapunov function W(x), given by G(V(x)) where V(x) is the optimum Liapunov function, i.e., a solution of equation (4.15). In order for W(x) to retain the basic nature of a Liapunov function, it will be required that G(V) satisfy the following conditions:

1) $G(V) > 0$ if $V > 0$

2) $G(0) = 0$

3) $dG(V)/dV > 0$ if $V > 0$

4) $\lim_{V \to \infty} G(V) = \infty$

5) $d^2G(V)/dV^2$ exists and is continuous.

The effect that this transformation has on the Hamilton-Jacobi equation can be observed by considering the total time derivative of W(x). Again $\dot{W}$ will be a function of both x and u and hence will be written $\dot{W}(x,u)$

$$\dot{W}(x,u) = \frac{dG(V)}{dV} \dot{V}(x,u) \tag{5.1}$$

Now minimizing $\dot{W}(x,u)$ with respect to all admissible controls, while remembering that V(x) and hence G(V(x)) is not a function of u, yields

$$\min_{u \in U} \dot{W}(x,u) = \frac{dG(V)}{dV} \min_{u \in U} \dot{V}(x,u)$$

$$= \frac{dG(V)}{dV} \dot{V}^o(x) \tag{5.2}$$

The minimum of $\dot{W}(x,u)$ with respect to $u \in U$ will be designated by $W^o(x)$. Then equation (5.2) becomes

$$\dot{W}^o(x) = \frac{dG(V)}{dV} \dot{V}^o(x) \tag{5.3}$$

But, by assumption, V(x) is a solution of equation (4.15) and hence $\dot{V}^{o}(x) = -L(x)$. Therefore equation (5.3) becomes

$$\dot{W}^{o}(x) = -L(x)\frac{dG(V)}{dV} \tag{5.4}$$

Since dG(V)/dV is positive for V greater than zero, G must be monotone increasing on the interval $(0,\infty)$. Then according to conditions 1) and 2) above, G must map the interval $[0,\infty)$ onto the interval $[0,\infty)$ in a one-to-one fashion. Therefore G possesses a unique inverse function I on the interval $[0,\infty)$. Since both V(x) and W(x) are required to be positive definite, this is the only region of interest. Therefore

$$V(x) = I(W(x)) \tag{5.5}$$

Then substituting for V(x) in equation (5.4) gives

$$\dot{W}^{o}(x) = -L(x)\frac{dG(I(W(x)))}{dV} \tag{5.6}$$

Now letting $F(W) = \frac{dG(I(W))}{dV}$, equation (5.6) becomes

$$\dot{W}^{o}(x) = -L(x)F(W(x)) \tag{5.7}$$

For the case of time optimal control, L(x) = 1, and equation (5.7) reduces to

$$\dot{W}^{o}(x) = -F(W(x)) \tag{5.8}$$

By combining the results of this section with the theorems of section 4.3, the following optimality theorem results.

Theorem 5.1 If there exists a Liapunov function, W(x), with continuous second partial derivatives, such that $\dot{W}^{o}(x) = -L(x)F(W(x))$ where F(W) = dG(I(W))/dV and G satisfies the conditions given above, then the control, u^{o}, which minimizes $\dot{W}(x,u)$, is an optimal control.

The first step in the proof of this theorem is to obtain the Liapunov function, V(x), which corresponds to W(x). Substituting W = G(V) into the definition of F(W) yields

$$F(G(V)) = dG(I(G(V)))/dV \tag{5.33}$$

However I is the inverse of G and hence I(G(V)) = V, then equation (5.33) becomes

$$F(G(V)) = dG(V)/dV \tag{5.34}$$

By antidifferentiation G(V) can be obtained from equation (5.34). By hypothesis this G(V) must satisfy the conditions given above. Hence V(x) given by I(W) must be a Liapunov function if W(x) is. Condition 5) on G(V) assures that if W(x) has continuous second partial derivatives that V(x) will also. Thus the first portion of Theorem 4.1 has been satisfied.

Next consider $\dot{V}(x,u)$ which may be obtained as

$$\dot{V}(x,u) = \frac{dV(x)}{dW} \dot{W}(x,u) \tag{5.35}$$

Since neither W(x) nor V(x) are functions of u, the same control u^o must minimize both $\dot{V}(x,u)$ and $\dot{W}(x,u)$ and hence equation (5.35) becomes

$$\dot{V}^o(x) = \frac{dV(x)}{dW} \dot{W}^o(x)$$

By hypothesis $\dot{W}^o(x) = -L(x)F(W(x))$ and therefore one obtains

$$\dot{V}^o(x) = -L(x) \left[dV/dW\right]\left[dG(V)/dV\right] \tag{5.36}$$

But G(V) = W and hence

$$dV/dW \left[dG(V)/dV\right] = \left[dV/dW\right]\left[dW/dV\right]$$
$$= 1.$$

Therefore equation (5.36) becomes

$$\dot{V}^{o}(x) = -L(x)$$

Hence V(x) satisfies the conditions of the Theorem 4.1 and u^{o} must be an optimum control which completes the proof of the theorem. For the minimum time problem, this theorem becomes

Theorem 5.2 If there exists a Liapunov function W(x), with continuous second partial derivatives, such that $\dot{W}^{o}(x) = F(W(x))$, then u^{o} is a time-optimal control.

This last theorem embodies the basic concept of the method presented by Nahi[14] for obtaining time-optimal control by the use of the Second Method. However, by the development presented here, greater insight and information are gained with regard to the function F.

It should be noted that since equation (5.7) is the modified Hamilton-Jacobi equation that $\nabla W(x)$ does not satisfy the adjoint equation even though $\dot{W}^{o}(x) = -L(x)F(W(x))$. This fact can readily be verified by example.

One transformation, G, which is of particular importance in the next section is

$$W = G(V) = \left[\frac{K_2}{K_1} \left(\exp \left(\frac{K_1}{2} V\right) - 1\right) \right]^2 \tag{5.9}$$

Then the inverse of G is given by

$$V = I(W) = \frac{2}{K_1} \ln\left(\frac{K_1\sqrt{W}}{K_2} + 1\right) \tag{5.10}$$

Therefore equation (5.7) becomes

$$\dot{W}^{o}(x) = -L(x) \left(K_1 W(x) + K_2 \sqrt{W(x)}\right) \tag{5.11}$$

For the time-optimal case, one obtains

$$\dot{W}^{o}(x) = -K_1 W(x) - K_2 \sqrt{W(x)} \tag{5.12}$$

This equation plays an important role in the next section.

5.3 Eigenvector Scalar Products

In this section a particular class of solutions of the Hamilton-Jacobi equation is developed. Because of the manner in which these solutions are formed, they are called eigenvector scalar product solutions. For the material to be presented in the remaining portion of this chapter and the next chapter, two additional assumptions are added to the basic optimization problem as formulated in section 2.3. First, only time-optimal control is considered, i.e., $L(x) = 1$. Second, the eigenvalues of the matrix $\underline{A}$ in equation (2.2) must be real, non-positive and distinct.

In the preceding section, it was shown that time optimal control could be obtained by finding a Liapunov function, $V(x)$, such that $\dot{V}^o(x) = -F(V(x))$. The following theorem, due to Malkin[10], establishes a necessary and sufficient condition for $\dot{V}(x) = \lambda V(x)$ for uncontrolled linear systems.

Theorem 5.3 For systems whose laws of motion are of the form $\dot{x} = \underline{A}x$ there exist Liapunov functions such that $\dot{V}(x) = \lambda V(x)$ if and only if $\lambda = m_1\lambda_1 + m_2\lambda_2 + \ldots + m_n\lambda_n$ and $V(x)$ is given by

$$V(x) = (q^{1'}x)^{m_1} (q^{2'}x)^{m_2} \ldots (q^{n'}x)^{m_n}$$

where the λ_i's are the eigenvalues of $\underline{A}$ and q_i is the eigenvector of $\underline{A}'$ associated with λ_i.

The reader is referred to Malkin[10] for a proof of the necessity portion of the above theorem, which is somewhat involved and not of particular importance for the present discussion. The proof of the sufficiency of the above theorem is presented below, since it is useful

in the following work. However, before beginning this proof, consider the following lemma.

Lemma 5.1 If q is an eigenvector of $\underline{A}'$ and λ is the associated eigenvalue, and if $V(x) = q'x$, then

$$\dot{V}(x) = \lambda V(x)$$

For $V(x) = q'x$ $\dot{V}(x)$ is given by

$$\dot{V}(x) = q'\dot{x} = q'\underline{A}x \tag{5.13}$$

But q is an eigenvector of A', hence

$$\underline{A}'q = \lambda q \tag{5.14}$$

Or, taking the transpose of both sides of equation (5.14), one obtains

$$q'\underline{A} = \lambda q' \tag{5.15}$$

Substituting equation (5.15) into equation (5.13) yields

$$\dot{V}(x) = \lambda q'x = \lambda V(x)$$

and the proof of the lemma is completed.

Returning to the proof of the theorem, consider a Liapunov function of the form

$$V(x) = (q^{1'}x)^{m1}\ (q^{2'}x)^{m2}\ \ldots\ (q^{n'}x)^{m_n} \tag{5.16}$$

Now let $V_i(x) = q^{i'}x$ and then equation (5.16) becomes

$$V(x) = V_1(x)^{m1}\ V_2(x)^{m2}\ \ldots\ V_n(x)^{m_n}$$

Now taking the total time derivative of V(x), one obtains

$$\dot{V}(x) = m_1 V_1^{m_1-1}\dot{V}_1 V_2^{m_2}\ldots V_n^{m_n} + \ldots + m_n V_1^{m_1}\ldots V_n^{m_n-1}\dot{V}_n \tag{5.17}$$

But from lemma 5.1, $\dot{V}_i = \lambda_i V_i$, then equation (5.17) becomes

$$\begin{aligned} V(x) &= \lambda_1 m_1 V^{m1} V_2^{m_2}\ldots\ V_n^{m_n} + \ldots + \lambda_n m_n V_1^{m_1}\ldots\ V_n^{m_n} \\ &= (\lambda_1 m_1 + \lambda_2 m_2 + \ldots + \lambda_n m_n)V(x) \end{aligned}$$

Thus completing the proof of the theorem.

Since $\dot{V}(x)$ turns out to be a function of $V(x)$, one is led to consider a Liapunov function of the form of equation (5.16) as a possible solution of the Hamilton-Jacobi or modified Hamilton-Jacobi equation. The following theorem indicates that there are, in fact, solutions of this form.

Theorem 5.4 If q is an eigenvector of $\underline{A}'$ and λ is the associated eigenvalue, then $W(x) = (q'x)^2$ is a solution of the modified Hamilton-Jacobi equation (5.12), i.e., $\dot{W}^o(x) = K_1 W(x) - K_2\sqrt{W(x)}$ where $-K_1 = 2\lambda$ and $K_2 = 2\|\underline{B}'q\|$.

As a first step in the proof, consider the following lemma.

Lemma 5.2 For any matrix $\underline{P}$ such that $\underline{P} = pp'$ and any matrix $\underline{B}$, $\underline{P}'\underline{B}\underline{B}'\underline{P} = \|\underline{B}'p\|^2\underline{P}$.

Writing out $\underline{P}'\underline{B}\underline{B}'\underline{P}$ in full, one obtains

$$\underline{P}'\underline{B}\underline{B}'\underline{P} = pp'\underline{B}\underline{B}'pp'$$

Now consider the $p'\underline{B}\underline{B}'p$ portion of this expression. $\underline{B}$ is an nxr matrix, while p is an nx1 column matrix (vector). Hence the product $p'\underline{B}$ is a 1 x r matrix, and $\underline{B}'p$ is an r x 1 matrix. Therefore the product $p'\underline{B}\underline{B}'p$ must be a 1 x 1 matrix, or a scalar, whose value is $\|\underline{B}'p\|^2$. Therefore

$$\begin{aligned}\underline{P}'\underline{B}\underline{B}'\underline{P} &= p(\|\underline{B}'p\|^2)p' \\ &= \|\underline{B}'p\|^2\underline{P}\end{aligned} \tag{5.18}$$

which completes the proof of the lemma. It should be pointed out that $\underline{B}$ is not required to be non-singular.

The next step, in the proof of the theorem, is to rewrite W(x) in a new form. Since $q'x = x'q$, then $W(x) = q'xq'x$ can be written as

$$\begin{aligned}W(x) &= x'qq'x \\ &= x'\underline{Q}x\end{aligned} \tag{5.19}$$

where $\underline{Q} = qq'$. It should be noted that $\underline{Q}$ is positive semidefinite and symmetric. Now taking the gradient of $W(x)$, one obtains

$$\nabla W(x) = 2\underline{Q}x \tag{5.20}$$

By substitution $W(x)$ for $V(x)$ in equation (4.5A), $\dot{W}^{o}(x)$ is given by

$$\dot{W}^{o}(x) = \nabla W'(x)\underline{A}x - \| \underline{B}'\nabla W(x)\| \tag{5.21}$$

Substituting equation (5.20) into equation (5.21) and expanding $\underline{B}'\nabla W(x^1)$, one obtains

$$\dot{W}^{o}(x) = 2x'\underline{Q}'\underline{A}x - 2\sqrt{x'\underline{Q}'\underline{B}\underline{B}'\underline{Q}x}$$

$$= 2x'qq'\underline{A}x - 2\sqrt{x'\underline{Q}'\underline{B}\underline{B}'\underline{Q}x} \tag{5.22}$$

But q is an eigenvector of $\underline{A}'$ and hence $q'\underline{A} = \lambda q'$. From the lemma above, $\underline{Q}'\underline{B}\underline{B}'\underline{Q} = \|\underline{B}'q\|^2\underline{Q}$. Therefore equation (5.22) becomes

$$\dot{W}^{o}(x) = 2\lambda(x'\underline{Q}x) - 2\|\underline{B}'q\|\sqrt{x'\underline{Q}x}$$

or

$$\dot{W}^{o}(x) = 2\lambda W(x) - 2\|\underline{B}'q\|\sqrt{W(x)} \tag{5.23}$$

Hence $W(x) = (q'x)^2$ satisfies equation (5.12) and the proof of the theorem is completed. Solutions of this type are called eigenvector scalar product solutions since they are scalar products of eigenvectors with the state vector.

By the use of equation (5.10), the Liapunov function, $V(x)$, which is a solution to the Hamilton-Jacobi equation (4.15) is given by

$$V(x) = \frac{1}{-\lambda} \ln \left(\frac{-\lambda|q'x|}{\|\underline{B}'q\|} + 1 \right) \tag{5.24}$$

It can be easily verified by direct substitution that $\dot{V}^o(x) = 1$. The corresponding optimal control is given by

$$u^o(x) = \frac{-\underline{B}'qq'x}{\|\underline{B}'qq'x\|} \tag{5.25}$$

The obvious simplicity of the form of W(x) as compared to V(x) points out the reason for the use of the modified Hamilton-Jacobi equation. However, V(x) is also important since surfaces of constant V(x) are surfaces of constant time. This point is discussed further in the next section, which is concerned with bounding the optimum transition time.

It should be pointed out that the solutions obtained above cannot be used directly since the Liapunov functions are only semidefinite. However, in the next chapter, a method of employing these solutions to obtain sub-optimal control is developed. Before proceeding to the next section, it is perhaps wise to consider a particular example of the solutions presented above.

Example 5.1 The equation of motion of the system are

$$\begin{bmatrix} \dot{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{5.26}$$

It is desired to find the eigenvector scalar product solutions for this problem and to show that they satisfy the modified Hamilton-Jacobi equation. The corresponding solutions of the Hamilton-Jacobi equation are also to be found and verified.

By standard methods the eigenvalues are found to be -1, -2 with the corresponding (unnormalized) eigenvectors of $\underline{A}'$ being (2,1) and

(1,1). It should be noted that any other set of eigenvectors of $\underline{A}'$ could have been chosen since the resulting optimal control and Liapunov function is unchanged. The above set was chosen for its computational convenience.

There are two solutions of the modified Hamilton-Jacobi equation which can be obtained by the above method, corresponding to the two eigenvectors.

First, for the eigenvalue -1, one obtains

$$W_1(x) = (q^{1'}x)^2 = (2x_1 + x_2)^2 \tag{5.27}$$

and $\dot{W}_1^o(x)$ is given by

$$\dot{W}_1^o(x) = -2W_1(x) - 2\sqrt{W_1(x)}$$

The corresponding solution of the Hamilton-Jacobi equation is

$$V_1(x) = \ln(|2x_1 + x_2| + 1) \tag{5.28}$$

while the optimum control as given by equation (5.25) is

$$u^o(x) = \frac{-(2x_1 + x_2)}{|2x_1 + x_2|} \tag{5.29}$$

The total time derivative of $V_1(x)$ is then given by

$$\dot{V}_1(x,u) = \left[\frac{1}{|2x_1 + x_2| + 1}\right]\left[\frac{(2x_1 + x_2)}{|2x_1 + x_2|}\right]\left[2\dot{x}_1 + \dot{x}_2\right]$$

Now substituting from equation (5.26) one obtains

$$\dot{V}_1(x,u) = \left[\frac{1}{|2x_1 + x_2| + 1}\right]\left[\frac{(2x_1 + x_2)}{|2x_1 + x_2|}\right]\left[-2x_1 - x_2 + u\right]$$

If $u^o(x)$ as given by equation (5.29) is now substituted for u, $\dot{V}_1(x,u)$ becomes $\dot{V}_1^o(x)$

$$\dot{V}_1^o(x) = \left[\frac{(2x_1 + x_2)}{|2x_1 + x_2|^2 + |2x_1 + x_2|}\right]\left[-(2x_1 + x_2) - \frac{(2x_1 + x_2)}{|2x_1 + x_2|}\right] = -1$$

Hence $V_1(x)$ satisfies the Hamilton-Jacobi equations as predicted. Then for the second eigenvalue, one obtains

$$W_2(x) = (q^{2'}x)^2 = (x_1 + x_2)^2 \tag{5.30}$$

and $V_2(x)$ is given by

$$V_2(x) = \frac{1}{2}\ln(2|x_1 + x_2| + 1) \tag{5.31}$$

Again it can be readily verified that $V_2(x)$ satisfies the Hamilton-Jacobi equation.

5.4 Bounds on Transition Time

In section 4.2, it was briefly mentioned that if $\dot{V}(x) = -L(x)$ then surfaces of constant $V(x)$ become surfaces of constant cost. This point perhaps needs further elaboration. In the case of time optimal control, $\dot{V}(x) = -1$, and hence integrating with respect to t from t_o to t_1, one obtains

$$V(x^1) - V(x^o) = t_o - t_1 \tag{5.32}$$

If the terminal state is taken to be the origin, then $V(x^1) = 0$, and

$$t_1 - t_o = V(x^o)$$

Thus the value of the Liapunov function at the initial state of the system is equal to the transition time. If a Liapunov function $V_o(x)$ has been found such that $\dot{V}_o{}^o(x) = -1$, then $V_o(x^o)$ is equal to the minimum transition time from x^o to the origin. Let S_o be the surface composed of all points x such that $V_o(x) = T$ where $V_o(x)$ is the solution of the Hamilton-Jacobi equation—the optimum Liapunov function. Then S_o is the set of all points from which it is possible to reach the origin in a transition time T_o by the use of time optimal control. This surface must be smooth and

enclose the origin. Figure 5.1 shows a two dimensional example where the surface S_o has become the closed curve designated by S_o. Such a surface will be called an isochrone. The problem of finding optimal control is actually a problem of finding the equation for the isochrone, or V(x).

Since it is normally impossible to obtain the exact solution of the Hamilton-Jacobi equation, it is necessary to approximate the solution. If such an approximate solution, $V_a(x)$, is found, then let S_1 be the surface composed of all points x such that $V_a(x) = T_o$, i.e., the set of all points from which the origin can be reached in T_o seconds by the use of sub-optimal control. The surface S_1 must be within or at most tangent to S_o as shown in Figure 5.1.

One method for judging the quality of a sub-optimal control is now obvious. The more nearly the surface S_1 coincides with the surface S_o, the better the sub-optimal control. However, since the surface S_o is generally not known, such a method of judging the quality of the approximation is rather academic. Some other method is therefore needed.

One such method is to find another surface S_2 which is entirely outside or at most tangent to S_o, as shown in Figure 5.1. If such a surface could be found in a relatively easy and straight-forward manner, the quality of an approximation could be determined in the following manner. If S_1 and S_2 were close, then S_1 must be a good approximation, since S_1 must be at least as close to S_o as it is to S_2. However, if S_1 and S_2 were far apart, no conclusion could be reached regarding the quality of the control since there would be no knowledge with respect to the relation of S_1 and S_o. This situation should be compared with the

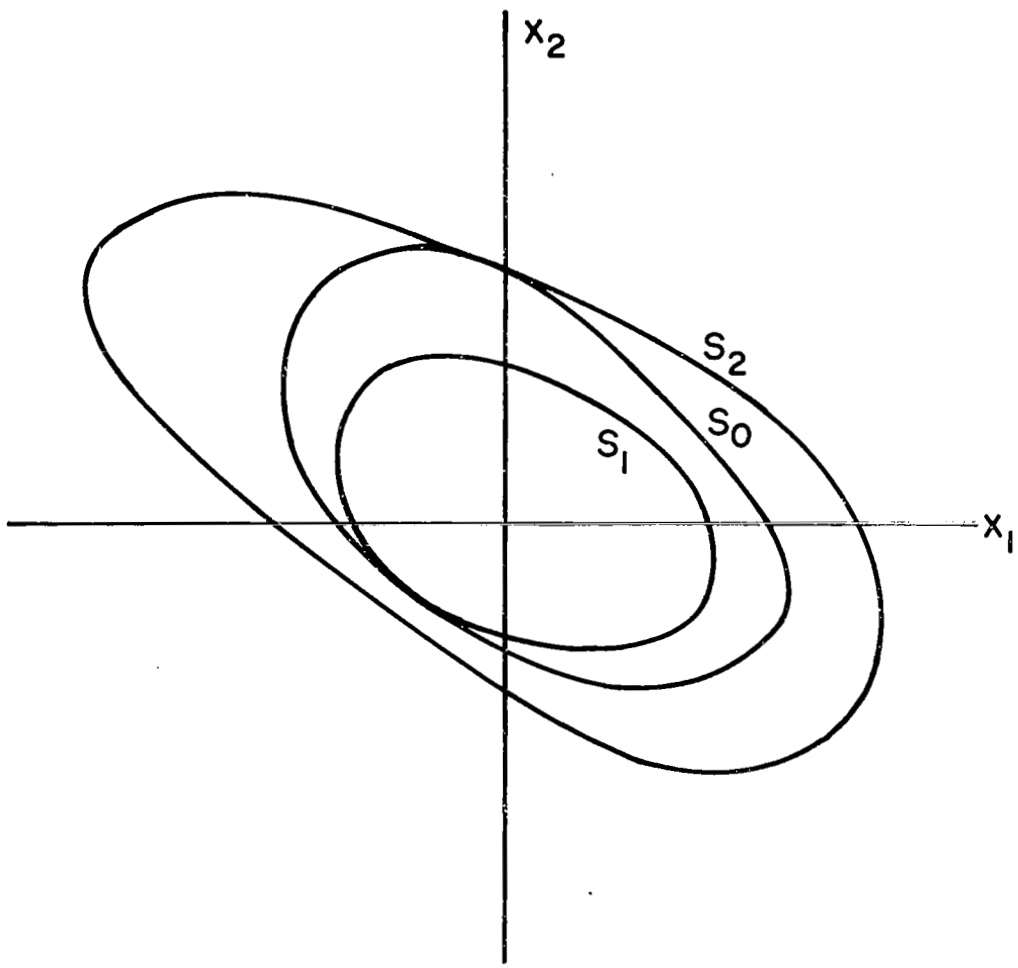


FIGURE 5.1 TYPICAL ISOCHRONES

basic concept of the Second Method where a failure to construct a Liapunov function generally yields no concrete results with respect to stability.

It should be noted that the surface S_2 does, in general, not correspond to any physical control situation. If there did exist an admissible control which would take the system from S_2 to the origin in T_o seconds, this would contradict the assumption that S_o was optimal. However, there may be points on S_2 which correspond to points on S_o, and hence from these points the system can be returned to the origin in T_o seconds.

The eigenvector scalar product solutions, as developed in the preceding section, provide an unusually simple method for obtaining a S_2-type surface. Although the surface generated does not uniformly approximate S_o from the outside, it is tangent to S_o at several points, as is pointed out later.

Consider for a moment the interpretation that one may give to Liapunov functions which are given by equation (5.24)

$$V(x) = \frac{1}{-\lambda} \ln \left(\frac{-\lambda |q'x|}{\|\underline{B}'q\|} + 1\right) \tag{5.24}$$

In this case $V(x^1)$ is zero if and only if $q'x^1$ is zero. Thus the value of $V(x^o)$ does not correspond to the minimum transition time from x^o to the origin but rather from x^o to the hyperplane defined by $q'x^1 = 0$. Since the surface $V(x) = T_o$ corresponds to the surface $|q'x| = K$, a constant, which is two hyperplanes, $V(x) = T_o$ is actually two hyperplanes symmetrically placed about $q'x = 0$. See Figure 5.2 for a two dimensional example of these $V(x)$ equals a constant surfaces and typical trajectories of the system.

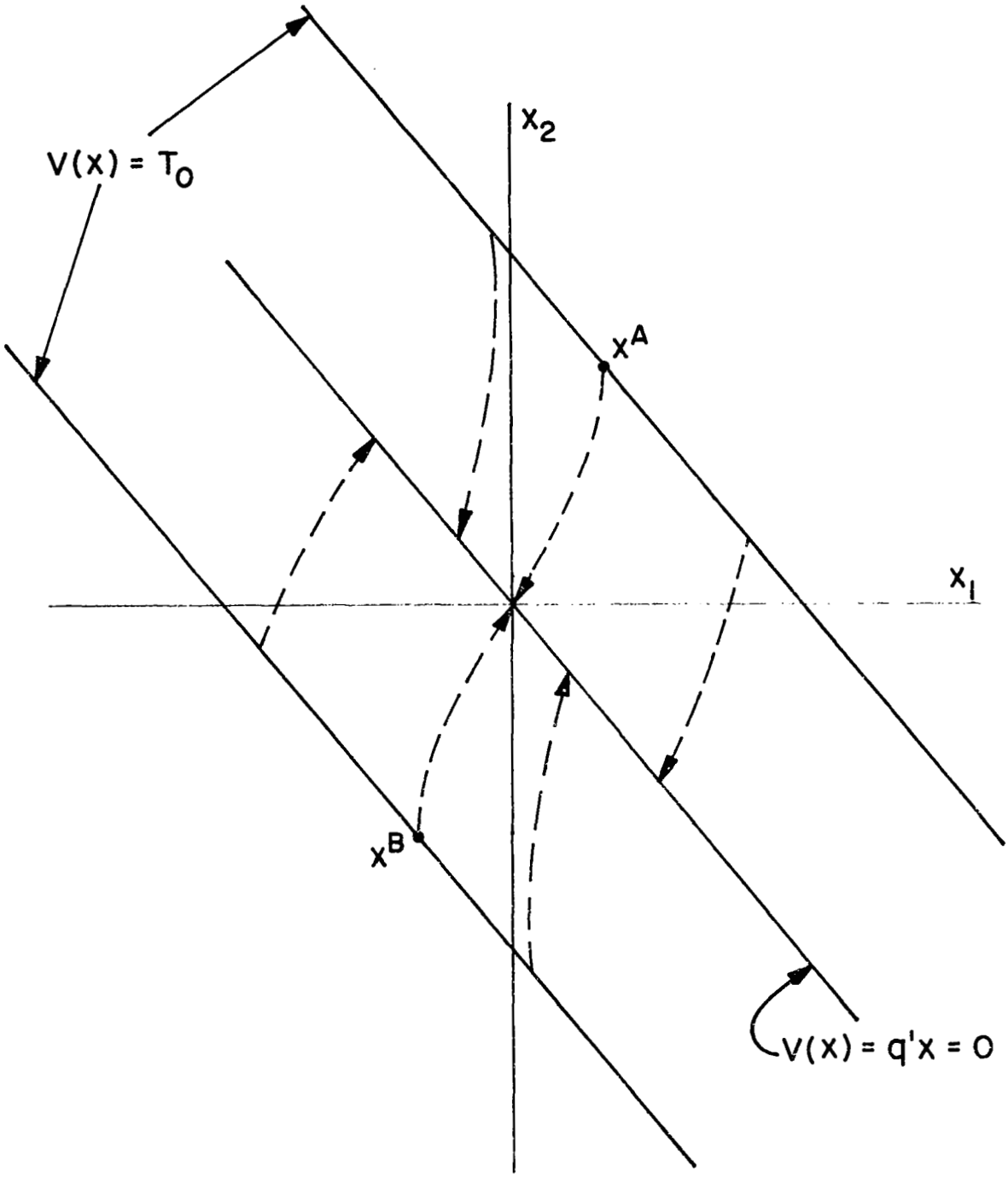


FIGURE 5.2 SURFACES OF CONSTANT V(x) FOR EIGENVECTOR SCALAR PRODUCT SOLUTIONS

Since the origin is one point on the hyperplane $V(x) = 0$, then $V(x^o)$ must be equal to or less than the minimum transition time from x^o to the origin. If it were greater, then there would exist a control which would transfer the system to the hyperplane in a time less than $V(x^o)$, which contradicts the optimality of the Liapunov function given by equation (5.24). Therefore the surface (hyperplanes) $V(x) = T_o$ must be entirely outside or at most tangent to the S_o surface.

It is very simple to show that the $V(x) = T_o$ surface must be tangent to S_o in two places. Since the system is controllable, there must be two points (one on each hyperplane) from which the origin is reached in T_o seconds as a special case of reaching the hyperplane $q'x = 0$. See points x^A and x^B in Figure 5.2. But these points must be on S_o; otherwise they would contradict the optimality of S_o. Hence there are two points at which the $V(x) = T_o$ surface is tangent to S_o.

Since the n eigenvalues are distinct, the eigenvectors are linearly independent and hence the n surfaces (hyperplanes) are non-coplanar (should probably be non-cohyperplanar). Therefore the boundary of the set of points for which $V_i(x) \leq T_o$, $i = 1,2,\ldots,n$ is a closed surface. See Figure 5.3 where the cross-hatched area is such a set. However every point on this surface must be outside or on the S_o surface, since each boundary point is on some surface $V_i(x) = T_o$, and by the argument above, each such point is outside or on S_o. Therefore this surface must be a S_2 surface.

The fact that there are 2n points at which the above S_2 surface is tangent to the S_2 surface can be argued in the following manner. By the argument presented above, the surface $V_1(x) = T_o$ must be tangent to

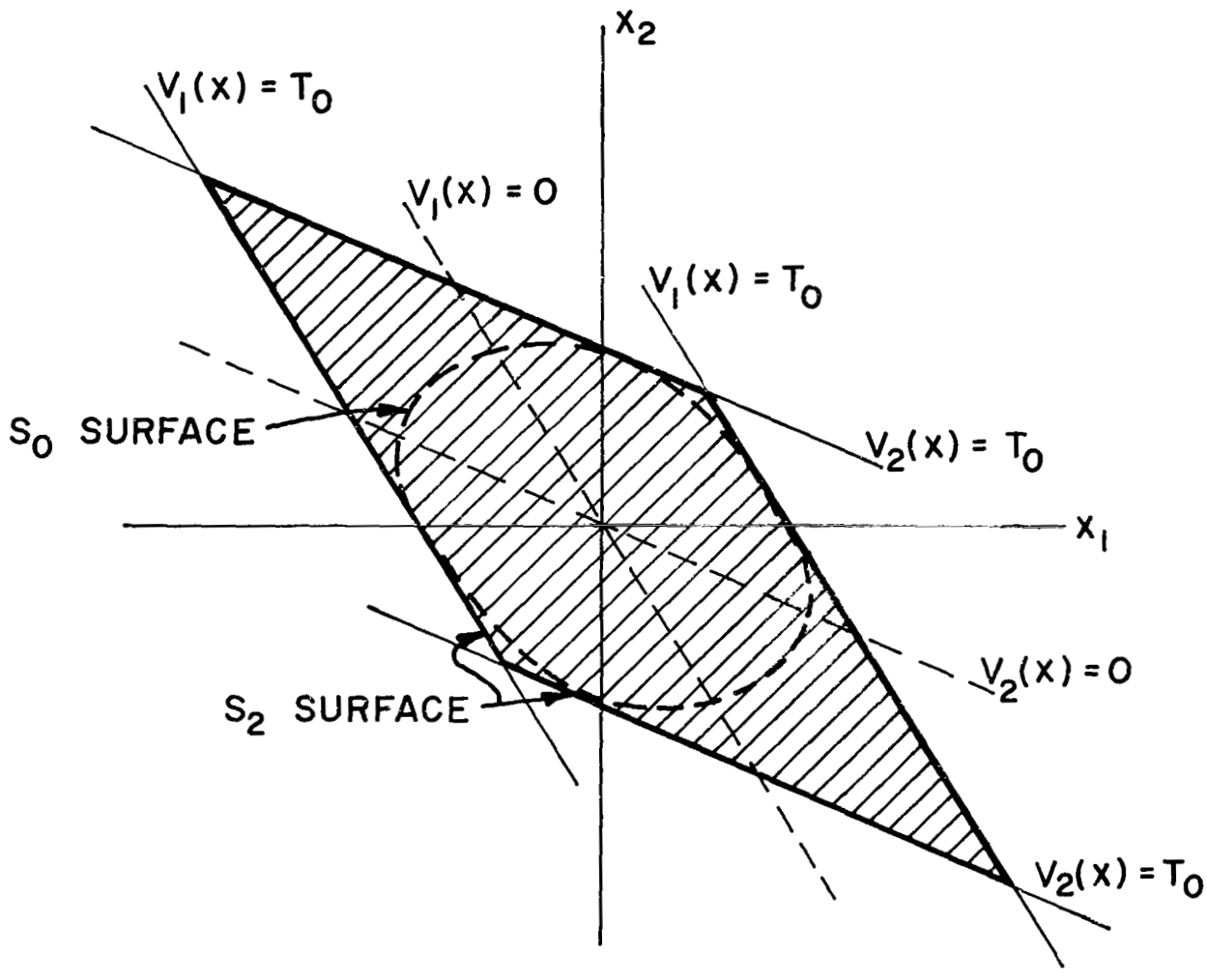


FIGURE 5.3 A TYPICAL S_2 SURFACE BOUNDING THE OPTIMAL S_0 SURFACE

S_o at two points. Since from these points it is possible to reach the origin in T_o seconds, it is also possible to reach all of the other $V_i(x) = 0$ surfaces in T_o seconds from these points. Therefore these points must be on both the S_2 and the S_o surfaces. By a similar argument, it can be concluded that there are 2n points which are common to S_2 and S_o. See Figure 5.3 which shows a typical S_2 surface generated by this method.

Thus by the use of the eigenvector scalar product solutions it is possible to obtain a relatively good S_2 surface in an unusually simple manner. The inability to find such surfaces has been a serious difficulty in designing approximately time-optimal systems in the past. Without such S_2 surfaces, it is impossible to judge the quality of a sub-optimal system without actually obtaining the optimal solution.

Chapter 6

SUB-OPTIMAL CONTROL FOR NON-SINGULAR B MATRIX

6.1 Introduction

In this chapter, a method for designing sub-optimal control systems is developed, based on the eigenvector scalar product solutions presented in the previous chapter. The control matrix, $\underline{B}$, is assumed to be non-singular for the work presented in this chapter. The method is developed first for second-order systems in order to be able to carry out a geometric representation and interpretation of the method. A second-order example completes the presentation.

Following the development of the sub-optimal control method for second-order systems, a generalization to n-th order systems is made. A third-order example is used to illustrate the generalization. The chapter concludes with a brief discussion of the method and its application.

It is perhaps of value to state the basic optimization problem that is considered in this chapter. For linear systems whose laws of motion are described by

$$\dot{x} = \underline{A}x + \underline{B}u$$

where the eigenvalues of $\underline{A}$ are real, distinct and non-positive and the matrix $\underline{B}$ is non-singular, it is desired to find time-optimal, closed-loop control corresponding to a transition from x^o to the origin. The control region, U, is the set of all control vectors, u, such that $\|u\|^2 \leq 1$.

6.2 Sub-Optimal Control of Second-Order Systems

In this section, a method of sub-optimal control of second-order systems with non-singular control matrices is developed. Before beginning this development, it is necessary to modify slightly the eigenvector scalar product solution of the previous chapter.

This modification involves a generalization of the bound on the norm of the control vector from unity to some unspecified constant, ρ. If such a change is made either by transforming the control vector or by repeating the work of Chapter 5, the Liapunov function, V(x), as given by equation (5.24) becomes

$$V(x) = -\frac{1}{\lambda} \ln \left(\frac{-\lambda |q'x|}{\rho \| \underline{B}'q \|} + 1\right) \tag{6.1}$$

and the corresponding optimal control is

$$u^o(x) = \frac{-\rho \underline{B}'qq'x}{\| \underline{B}'qq'x \|} \tag{6.2}$$

As would be expected, for a fixed initial condition, x^o, increasing ρ causes $V(x^o)$ to decrease, i.e., the transition time decreases with increasing control effort. Since the numerical value of V(x) is dependent on both the system's state and on the norm of the control vector, V(x) will be written as $V(x,\rho)$ to indicate this relation. Similarly, u will be written as $u(x,\rho)$.

A general second-order system with real, non-positive and distinct eigenvalues, λ_1 and λ_2 is considered in this section. For each eigenvalue, there is an eigenvector, designated by q^1 and q^2 respectively. Associated with the two eigenvectors are two Liapunov functions given by

equation (6.1), $V_1(x,\rho)$ and $V_2(x,\rho)$, and their corresponding optimal controls $u^1(x,\rho)$ and $u^2(x,\rho)$.

For some point in the state space, x^o, the control given $u^1(x,\rho)$ transfers the system from x^o to some point, x^1, on the surface $V_1(x,\rho) = 0$ in minimum time. Typical points and the corresponding optimal trajectory are shown in Figure 6.1. In the case of second-order systems, surfaces of V(x) equals a constant become lines. That $u^1(x,\rho)$ is, in fact, constant during this transition can be shown in the following manner.

For a given eigenvector, q^1, $\underline{B}'q^1$ in equation (6.2) is a constant vector while $q^{1\prime}x$ is a scalar. Hence $u^1(x,\rho)$ must be a vector parallel to $\underline{B}'q^1$; its direction being determined by the sign of the scalar quantity $q^{1\prime}x$. Since the $x_i(t)$'s are continuous functions of time, it is necessary for $q^{1\prime}x$ to be zero before it can change sign. But if $q^{1\prime}x$ is zero, then $V_1(x,\rho)$ is also zero. Therefore the sign of $q^{1\prime}x$ cannot change during the transition from x^o to x^1. Hence $u^1(x,\rho)$ must be a constant vector, whose norm is equal to ρ and whose direction is given by $\underline{B}'q^1q^{1\prime}x$. Figure 6.2 shows a typical control vector, $u^1(x,\rho)$.

Consider now another constant control vector, u, as shown in Figure 6.2 which is equal to the addition of $u^1(x,\rho)$ and any arbitrary constant vector r perpendicular to $u^1(x,\rho)$. Therefore

$$u = u^1(x,\rho) + r$$

where r is any constant vector such that $r'u^1(x,\rho) = 0$. The transition time from the point x^o to the line $V_1(x,\rho) = 0$ is independent of r; this

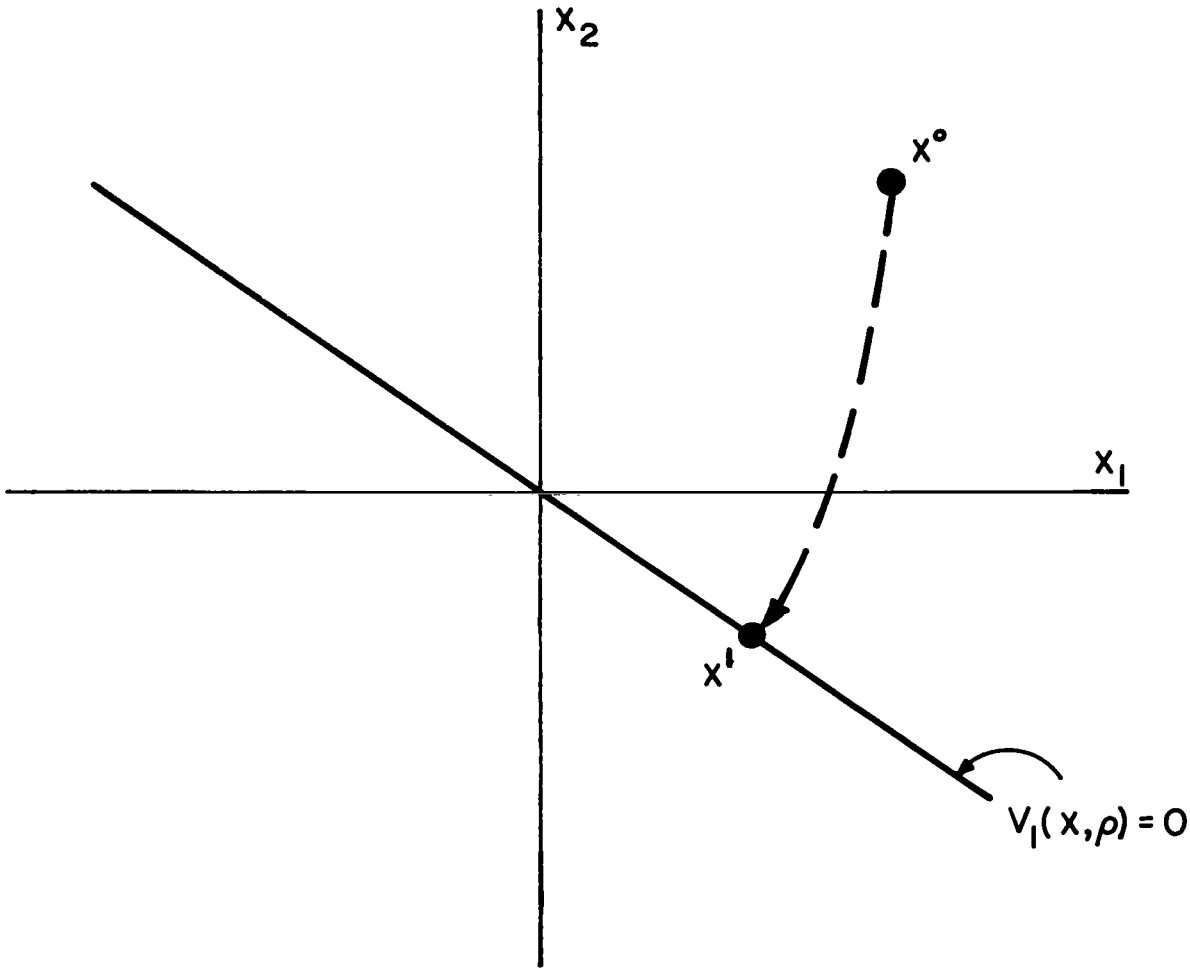


FIGURE 6.1 TYPICAL EIGENVECTOR SCALAR PRODUCT OPTIMAL TRAJECTORY

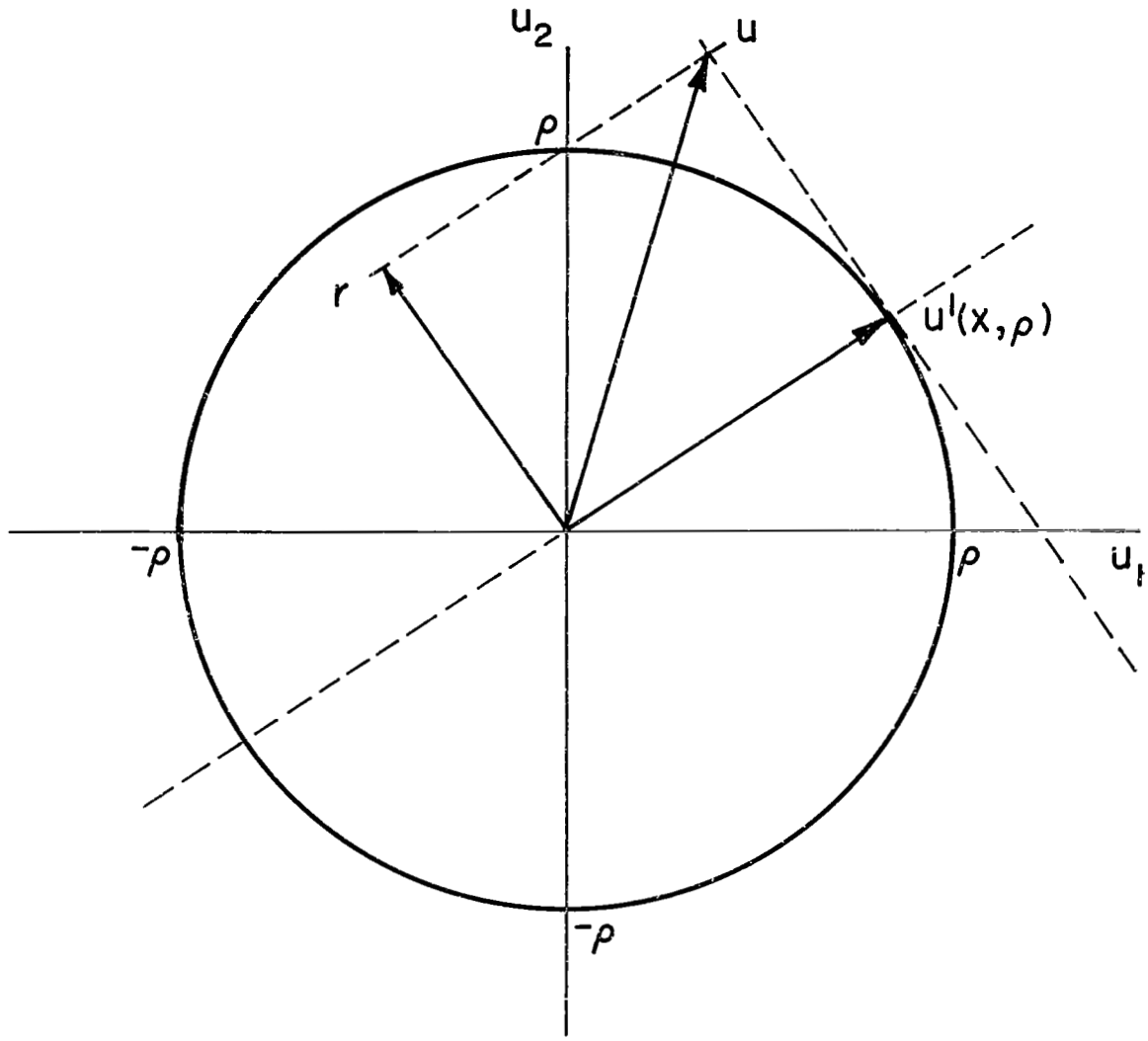


FIGURE 6.2 TYPICAL OPTIMAL AND NON-OPTIMAL CONTROLS

fact can be shown in the following manner.

Consider the Liapunov function $V_1(x,\rho)$ as given by equation (6.1). Now computing its total time derivative, one obtains

$$\dot{V}_1(x,\rho) = \frac{x'qq'\dot{x}}{-\lambda|q'x|^2 + \rho\|\underline{B}'qq'x\|}$$

$$= \frac{x'qq'(\underline{A}x + \underline{B}u)}{-\lambda|q'x|^2 + \rho\|\underline{B}'qq'x\|} \tag{6.4}$$

Substituting equation (6.3) for u gives

$$\dot{V}_1(x,\rho) = \frac{x'qq'\underline{A}x + x'qq'\underline{B}u^1(x,\rho) + x'qq'\underline{B}r}{-\lambda|q'x|^2 + \rho\|\underline{B}'qq'x\|} \tag{6.5}$$

However, r is perpendicular to $u^1(x,\rho)$ and by the argument above $u^1(x,\rho)$ is parallel to $\underline{B}'q$. Therefore r must be perpendicular to $\underline{B}'q$ and the scalar product of r and $\underline{B}'q$ must be zero, i.e., $q'\underline{B}r = 0$. Hence the third term in the numerator of equation (6.5) must be zero. Therefore

$$\dot{V}_1(x,\rho) = \frac{x'qq'\underline{A}x + x'qq'\underline{B}u^1(x,\rho)}{-\lambda|q'x|^2 + \rho\|\underline{B}'qq'x\|} \tag{6.6}$$

By direct substitution of $u^1(x,\rho)$ as given by equation (6.2), it can be readily verified that $\dot{V}_1(x,\rho) = -1$. Since neither $V_1(x,\rho)$ nor $\dot{V}_1(x,\rho)$ are functions of r it is obvious that the transition time from x^o to the line $V_1(x,\rho) = 0$ is independent of r.

From the argument above, one may conclude that for any control u only that portion of u which is parallel to $u^1(x,\rho)$ is important in determining the transition time from an initial point to the line $V_1(x,\rho) = 0$. One may draw a similar conclusion for u^2 and V_2.

For some initial state, x^o, let the magnitude of the optimum control vectors, $u^1(x^o,\rho_1)$ and $u^2(x^o,\rho_2)$, be chosen such that $V_1(x^o,\rho_1) = V_2(x^o,\rho_2) = T$. Therefore the control vector $u^i(x^o,\rho_i)$ transfers the system from x^o to the line $V_i = 0$ in T seconds. Now consider a vector u such that the portion of u which is parallel to u^i is equal to $u^i(x^o,\rho_i)$. See Figure 6.3 for a graphical representation of this situation. The magnitude, ρ_i, of the portion of u which is parallel to u^i may be obtained from the scalar product of u and a unit vector parallel to u^i. Hence

$$\rho_i = u'u^i(x^o,1) \tag{6.7}$$

Since only the portion of u which is parallel to u^i has any effect on the time necessary to transfer the system from x^o to $V_i = 0$, u must transfer the system from x^o to both $V_1 = 0$ and $V_2 = 0$ in the same time. But $V_1 = V_2 = 0$ can only occur at the origin, hence u must transfer the system from x^o to the origin in $V_1(x^o,\rho_1) = V_2(x^o,\rho_2) = T$ seconds.

Since setting $V_1(x^o,\rho_1) = V_2(x^o,\rho_2)$ specifies only the relative magnitude of ρ_1 in terms of ρ_2, there is an infinite number of vectors which satisfy this condition. However, only one of these vectors has unit length. This is then an admissible control which transfers the system from x^o to the origin in a finite time, $V_1(x^o,\rho_1)$. This is, in general, not the minimum time, but it is an acceptable compromise between system complexity and speed of response, as is shown later.

Several significant aspects of this sub-optimal control method should be noted. First, once the control is obtained, it is constant for the entire transition time. For small disturbances, the control

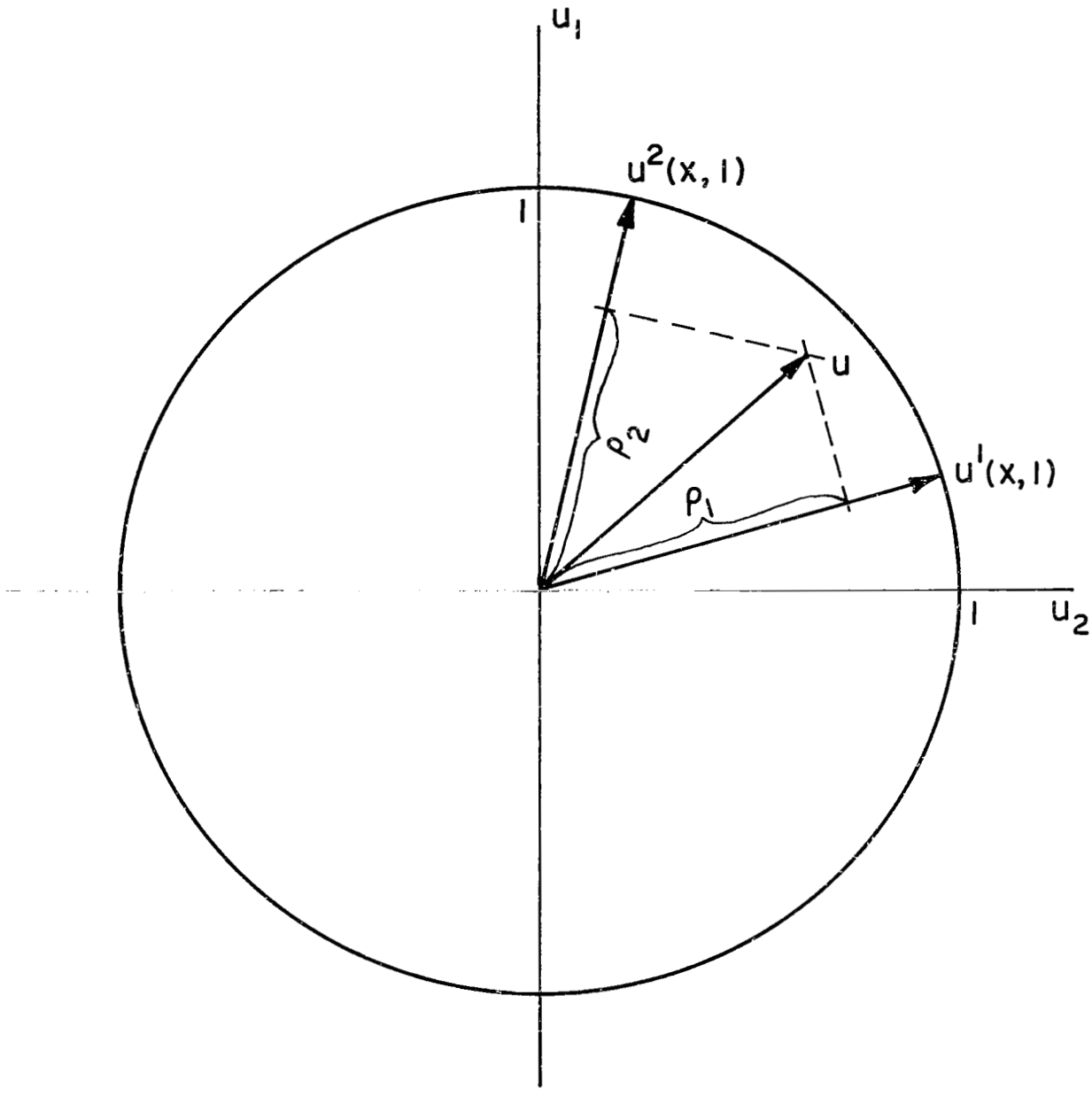


FIGURE 6.3 REPRESENTATION OF A VECTOR IN TERMS OF THE OPTIMAL CONTROL VECTORS

varies only slightly, which is helpful in mechanizing the controller. The control vector does not require rapid variations after its initial setting and hence only a minimum of recalculation of the control vector is necessary during the transition time. This should simplify the instrumentation of the controller.

Second, by the use of this method the transition time from any point to the origin can be easily obtained. This may be done by first setting the norm of the control vector equal to unity and then solving $V_1(x^o,\rho_1) = V_2(x^o,\rho_2) = T$ for T, which is the desired transition time. Isochrones can be found by choosing a value of T and finding the set of all points x such that $V_1(x,\rho_1) = V_2(x,\rho_2) = T$ and $\|u\| = 1$.

An interesting aspect of such isochrones is that for a given T, they are quadratic in terms of x_1 and x_2. This fact may be useful in implementation of the sub-optimal control method.

This method also makes it possible to obtain easily and directly the actual trajectory of the system from x^o to the origin. This can be done in the following manner. After finding the transition time, T_o, as described above, choose any time $T < T_o$; then solve for the point x such that $V_1(x,\rho_1) = V_2(x,\rho_2) = T$ with ρ_1 and ρ_2 as given above. This is the state of the system at T seconds before reaching the origin, or $T_o - T$ seconds after leaving the initial state. This allows one to obtain the position of the system at any time during the transition of the origin with no knowledge of any previous state, thus eliminating any accumulation of error. The work involved is strictly of an algebraic nature; it is not necessary to solve any differential equation. Both of

these last two aspects of the method aid one in evaluating whether the performance of the sub-optimal system is satisfactory.

One further aspect of this method should be mentioned because of its importance relative to the implementation of the method. The simultaneous solution of the equations $V_1(x^o,\rho_1) = V_2(x^o,\rho_2)$ and $\|u\| = 1$ is an algebraic manipulation, although it is not trivial. This should be contrasted to many of the presently advocated methods for which it is necessary to solve simultaneously the usual nonlinear differential equations of the two point boundary value nature. The computational advantage is obvious from a hardware standpoint. Since these computations are algebraic, it is possible to carry them out continuously on an analog computer to create continuous control.

Before considering a numerical example to illustrate the method, it is perhaps of value to outline the complete method for reference.

1) Obtain the eigenvalues and eigenvectors of the matrix $\underline{A}'$.

2) Obtain the two Liapunov functions as given by equation (6.1), $V_1(x,\rho)$ and $V_2(x,\rho)$, and their corresponding optimal controls, $u^1(x,\rho)$ and $u^2(x,\rho)$.

3) For a given point, x^o, solve the relations $\rho_i = u'u^i(x^o,1)$ to obtain u_1 and u_2 in terms of ρ_1 and ρ_2.

4) Solve $V_1(x^o,\rho_1) = V_2(x^o,\rho_2)$ and $\|u\| = 1$ simultaneously to obtain ρ_1 and ρ_2.

5) By the use of the relations obtained in step three, find u, the desired sub-optimal control.

A method of mechanizing the last three steps of this procedure by the use of a digital or analog computer to create a closed-loop system is shown schematically in Figure 6.4. Two points should be emphasized again. First, once the control is determined, it remains relatively

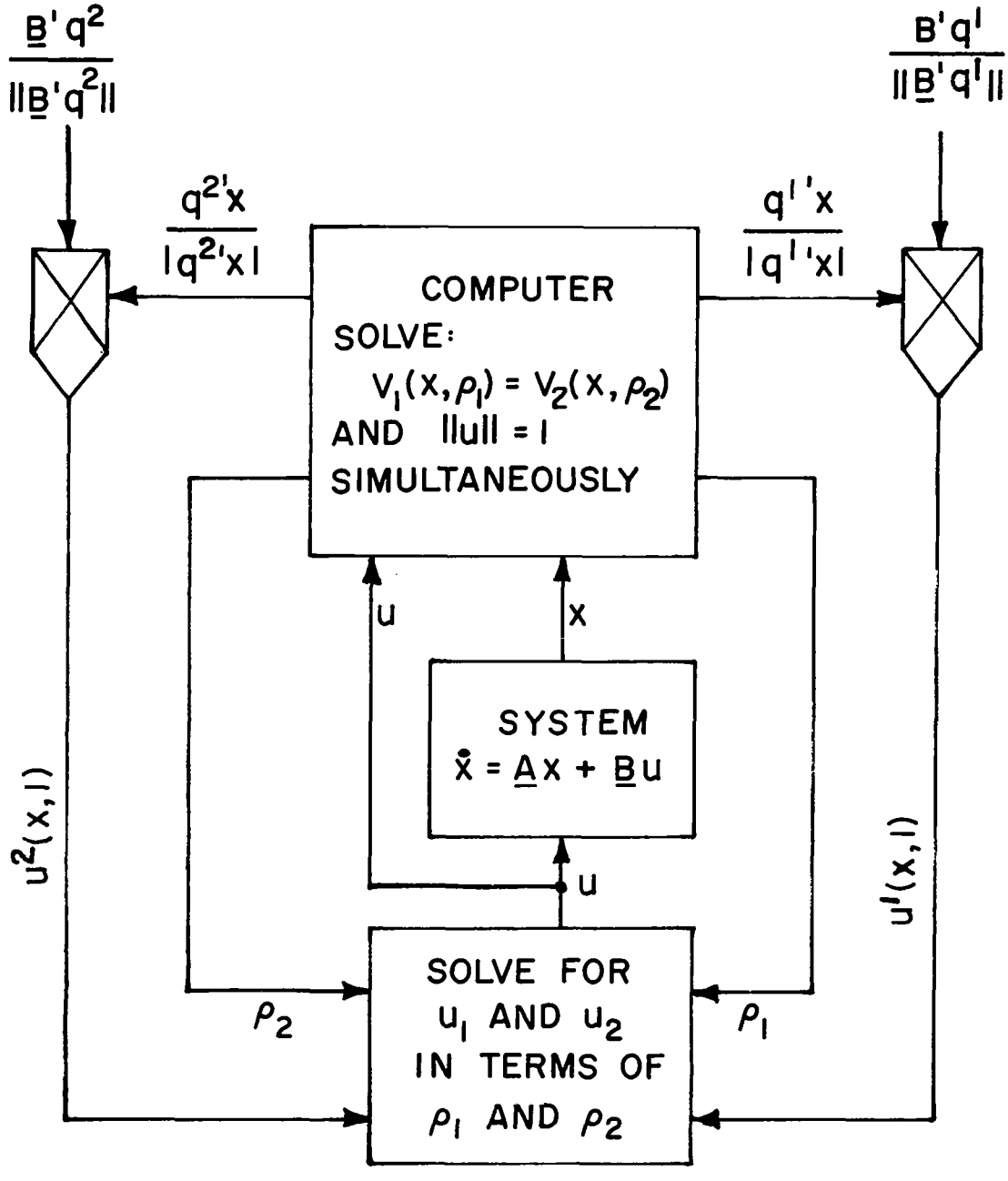


FIGURE 6.4 CLOSED-LOOP, SUB-OPTIMAL CONTROL SYSTEM

constant. Second, the operations required of the computer are strictly algebraic. It should also be noted that although the procedure is given in a step-by-step fashion, the control can be computed continuously by the use of analog computer.

Example 6.1 In order to illustrate the method of sub-optimal control developed above, consider the following system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & -1/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{6.9}$$

It is desired to transfer the system from the point $x^o = (2,1)$ to the origin with $u = 1$.

By standard methods the eigenvalues are found to be -1 and -2 with the corresponding eigenvectors of $\underline{A}'$ being (2,1) and (1,1). The two Liapunov functions as given by equation (6.1) are

$$V_1(x,\rho) = \ln \left(\frac{|2x_1 + x_2|}{2\rho} + 1\right) \tag{6.10}$$

$$V_2(x,\rho) = \frac{1}{2} \ln \left(\frac{4|x_1 + x_2|}{\sqrt{5}\,\rho} + 1\right) \tag{6.11}$$

The corresponding optimal controls as given by equation (6.2) are

$$\begin{bmatrix} u_1^1(x,\rho) \\ u_2^2(x,\rho) \end{bmatrix} = -\rho \left(\frac{2x_1 + x_2}{|2x_1 + x_2|}\right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{6.12}$$

$$\begin{bmatrix} u_1^2(x,\rho) \\ u_2^2(x,\rho) \end{bmatrix} = -\rho \left(\frac{x_1 + x_2}{|x_1 + x_2|}\right) \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \tag{6.13}$$

This completes steps one and two in the procedure outlined above.

Now for $x^o = (2,1)$, $u^1(x^o, 1)$ and $u^2(x^o, 1)$ become

$$u^1(x^o, 1) = (-1,0) \tag{6.14}$$

$$u^2(x^o, 1) = (-2/\sqrt{5}, -2/\sqrt{5}) \tag{6.15}$$

Using the relation $\rho_i = u^t u^1(x^o, 1)$, one obtains

$$\rho_1 = -u_1$$

$$\rho_2 = -2u_1/\sqrt{5} \quad -u_2/\sqrt{5}$$

Then solving for u_1 and u_2 in terms of ρ_1 and ρ_2 yields

$$u_1 = -\rho_1$$

$$u_2 = +2\rho_1 - \sqrt{5}\ \rho_2$$

Now setting $\|u\|^2 = 1$, one obtains

$$(-\rho_1)^2 + (-2\rho_1 + \sqrt{5}\ \rho_2)^2 = 1 \tag{6.16}$$

Setting $V_1(x^o,\rho_1) = V_2(x^o,\rho_2)$, one obtains

$$\ln\left(\frac{5}{2\rho_1} + 1\right) = 1/2 \ln\left(\frac{12}{\sqrt{5}\rho_2} + 1\right) \tag{6.17}$$

By solving equations (6.16) and (6.17) simultaneously, it is possible to obtain ρ_1 and ρ_2. If this is done the solution obtained is

$$\rho_1 = 0.645$$

$$\rho_2 = 0.236$$

Therefore a control vector $u = (-0.645, +0.763)$ transfers the system from the point $x^o = (2,1)$ to the origin in $V_1(x^o,\rho_1) = 1.58$ seconds.

6.3 Generalization

In the previous section a sub-optimal method was developed for second-order systems. A generalization of this method to n-th order systems is presented in this section. Since all of the proofs and arguments carry over directly to the n-th order case, only the conclusions are presented here.

In the n-th order case, there are n real and distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and hence n linearly independent eigenvectors $q^1, q^2, \ldots, q^n$. Associated with each eigenvector is a Liapunov function given by equation (6.1) $V_1(x,\rho), V_2(x,\rho), \ldots, V_n(x,\rho)$ and their corresponding optimal controls $u^1(x,\rho), u^2(x,\rho), \ldots, u^n(x,\rho)$. As before, for some point in the state space, x^o, the control given by $u^i(x,\rho)$ transfers the system from x^o to some point, x^1, on the hyperplane $V_1(x,\rho) = 0$ in minimum time. Again the control $u^i(x,\rho)$ is constant during the entire transition time.

By an argument identical to that presented in the previous section, it can be shown that for any control u only that portion of u which is parallel to $u^i(x,\rho)$ affects the transition time from any initial point to the hyperplane $V_i(x,\rho) = 0$.

If for some initial state x^o a control vector u is chosen such that $V_1(x^o,\rho_1) = V_2(x^o,\rho_2) = \ldots = V_n(x^o,\rho_n)$ where $\rho = u'u^i(x,1)$, then this control must transfer the system to the origin in $V_1(x^o,\rho)$ seconds.

Since there are only n-1 equations in the n unknowns, $\rho_1, \rho_2, \ldots, \rho_n$, there is an infinite set of control vectors which satisfy these equations. From this set, there is only one vector whose norm is equal to unity. This is an admissible control which transfers the system from any point x^o to the origin in a finite, although usually not minimum, time.

Then the procedure for obtaining a sub-optimal control can be stated in the following steps:

1) Obtain the eigenvalues and eigenvectors of the matrix $\underline{A}'$.

2) Obtain the Liapunov functions as given by equation (6.1), $V_1(x,\rho)$ $V_2(x,\rho)$, ..., $V_n(x,\rho)$ and their corresponding optimal controls, $u^1(x,\rho)$, $u^2(x,\rho)$, ..., $u^n(x,\rho)$.

3) For a given point, x^o, solve the relations $\rho_i = u'u^i(x^o,1)$ to obtain u_i in terms of the ρ_i.

4) Solve $V_1(x^o,\rho_1) = V_2(x^o,\rho_2) = \ldots = V_n(x^o,\rho_n)$ and $\|u\| = 1$ simultaneously to obtain $\rho_1, \rho_2, \ldots, \rho_n$.

5) By the use of the relations obtained in step three, find u, the desired sub-optimal control.

As before the last three steps in this procedure can be mechanized by the use of a digital or analog computer in order to create a closed-loop system.

All of the features of this method which were pointed out for second-order systems carry over directly for n-th order systems.

Example 6.2 As an illustration of the above procedure, consider the third-order system shown in Figure 6.5. The equations of motion may be written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -3 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

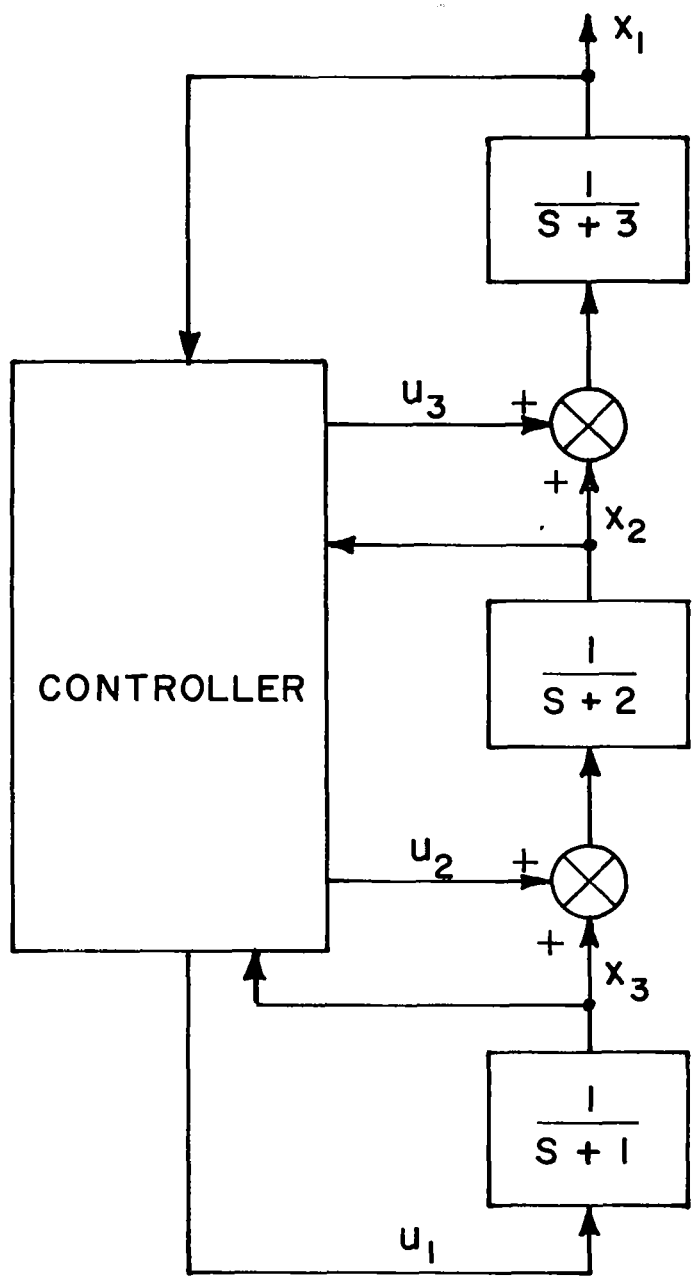


FIGURE 6.5 THIRD ORDER EXAMPLE FOR SUB-OPTIMAL CONTROL WITH NON-SINGULAR CONTROL MATRIX

It is desired to find a sub-optimal control which transfers the system from the point $x^o = (1,2,3)$ to the origin. The norm of the control vector is constrained to be equal to or less than unity.

The eigenvalues are -1, -2, and -3 with the corresponding eigenvectors of $\underline{A}'$ being (0,0,1), (0,1,-1) and (2,-2,1). The Liapunov functions are found from equation (6.1) to be

$$V_1(x,\rho) = \ln\left(\frac{|x_3|}{\rho} + 1\right)$$

$$V_2(x,\rho) = \frac{1}{2}\ln\left(\frac{\sqrt{2}|x_2 - x_3|}{\rho} + 1\right)$$

$$V_3(x,\rho) = \frac{1}{3}\ln\left(\frac{|2x_1 - 2x_2 + x_3|}{\rho} + 1\right)$$

The corresponding optimal controls are then

$$\begin{bmatrix} u_1^1(x^o,1) \\ u_2^1(x^o,1) \\ u_3^1(x^o,1) \end{bmatrix} = -\left(\frac{x_3}{|x_3|}\right) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u_1^2(x^o,1) \\ u_2^2(x^o,1) \\ u_3^2(x^o,1) \end{bmatrix} = -\left(\frac{x_2 - x_3}{|x_2 - x_3|}\right) \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} u_1^3(x^o,1) \\ u_2^3(x^o,1) \\ u_3^3(x^o,1) \end{bmatrix} = -\left(\frac{2x_1 - 2x_2 + x_3}{|2x_1 - 2x_2 + x_3|}\right) \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix}$$

By the use of the initial state $x^o = (1,2,3)$ and the relations $\rho_i = u'u^i(x^o,1)$, it is possible to solve for the components of the control vector in terms of the ρ_i's.

$$u_1 = -\rho_1/2 + \sqrt{2}\,\rho_2 - 3\rho_3/2$$

$$u_2 = -\rho_1 + \sqrt{2}\,\rho_2$$

$$u_3 = -\rho_1$$

By setting $V_1(x^o,\rho_1) = V_2(x^o,\rho_2) = V_3(x^o,\rho_3)$, one obtains the following two equations

$$\ln\left(\frac{3}{\rho_1} + 1\right) = 1/2\,\ln\left(\frac{\sqrt{2}}{\rho_2} + 1\right) \tag{6.18}$$

$$\ln\left(\frac{3}{\rho_1} + 1\right) = 1/3\,\ln\left(\frac{1}{\rho_3} + 1\right) \tag{6.19}$$

In order to obtain the desired control vector, it is necessary to solve equations (6.18) and (6.19) simultaneously with $\|u\| = 1$. The answers that one obtains are

$$\rho_1 = 0.714$$

$$\rho_2 = 0.0543$$

$$\rho_3 = 0.00716$$

The desired sub-optimal control is then

$$u(x^o) = (-0.291, -0.637, -0.714)$$

This control transfer the system from $x^o = (1,2,3)$ to the origin in 1.65 seconds.

6.4 Discussion of the Method

In the previous sections, a method of obtaining sub-optimal control of systems in which the control matrix is non-singular was developed. Every sub-optimal control method should satisfy two requirements. First, the method should make it easy to design and implement the sub-optimal control system. Second, the performance of the sub-optimal control system should be acceptably close to the true optimum.

This method has several aspects which assist in the design and implementation of the sub-optimal system. These points have been discussed in section 6.2, but are repeated here for reference. In the absence of a disturbance the control vector, once obtained, remains constant until the system reaches the origin. The transition time from any point to the origin as well as the trajectory to the origin can be obtained readily. The isochrones are easy to find. In designing a closed-loop control system using this method, it is necessary for the controller-computer to solve only algebraic equations thus allowing continuous control.

Until now, the quality of the performance of the sub-optimal has been ignored. It is shown in this section that the quality is acceptable. Because of the difficulty involved, it is not possible to obtain the true optimal solution and hence it is necessary to use the approach discussed in section 5.4. In particular, it is shown that the sub-optimal isochrone, S_1, is tangent to the optimum isochrone, S_o, at several points.

As was pointed out in section 5.4, there must be two points on the $V_1(x,1) = T_o$ surface from which the origin is reached in T_o seconds

as a special case of reaching the surface using the control $u = u^1(x,1)$. Since the system reaches the origin in T_o seconds, it must also reach all of the $V_i(x,1) = 0$ surfaces in the same time. Hence $V_1(x^o,1) = V_2(x^o,\rho_2) = \ldots = V_n(x^o,\rho_n)$, where $\rho_i = u^1(x,1)'u^i(x^o,1)$, is satisfied at this point. The control $u^1(x^o,1)$ therefore satisfies all of the conditions of the sub-optimal control and hence it is the sub-optimal for these points. Therefore these points must be on the sub-optimal isochrone. But it is shown in section 5.4 that these points are also on the S_o surface. Hence the sub-optimal and optimal isochrones must be tangent at these points.

In a similar fashion, one could argue that there are two points on each $V_i(x,1) = T_o$ surface which are on both the optimal and sub-optimal isochrones. Hence there must be 2n points at which these surfaces are tangent. Since both of the surfaces are smooth, it is logical to assume that they are close in some region about each of these points.

One could get a direct measure of the quality of the sub-optimal control by determining the optimal isochrones for particular problems such as the ones in examples 6.1 and 6.2. However, the advisability of this is highly questionable. First, as was pointed out in Chapter 2, the computational labor involved in obtaining the optimum solution for even one point is horrendous for all but trivial problems. To find a complete set of such points is almost unthinkable. Second, if one were to carry out such computations, the most that one could conclude would be that the sub-optimal method was good or bad for that particular example.

It appears reasonable from the above points to conclude that this sub-optimal control method represents an acceptable compromise

between system complexity and speed of response.

Although the method presented in this chapter is significant and important in its own context, its major significance is in providing an underlying framework for future research. In the case of non-singular $\underline{B}$ matrix, several other sub-optimal methods have been suggested. None of these methods, however, have, as of yet, produced a sub-optimal control better than that presented here. The number of practical systems for which $\underline{B}$ is non-singular is limited and hence additional research is needed to extend this method to the case where the $\underline{B}$ matrix is singular.

Chapter 7

CONCLUSIONS AND FUTURE RESEARCH TOPICS

7.1 Conclusions

In this work, the Second Method of Liapunov was combined with the minimum principle to form a basis for a method of closed-loop, approximately time-optimal control of linear systems with bounded control of linear systems with bounded control norm. The first step was to show that solving the basic optimization problem is equivalent to solving a first-order partial differential equation which is identical to the Hamilton-Jacobi equation. Although it was not possible to solve this equation is general, a special class of solutions was shown to exist which provide a foundation for a sub-optimal control method.

By the use of these solutions, called eigenvector scalar products, it is possible to find surfaces which bound the optimal isochrones from the outside. The inability to find such surfaces has been a serious difficulty in designing approximately time-optimal systems in the past. Although this surface does not approximate the optimal isochrone uniformly it is shown that the two surfaces are tangent at 2n points.

The eigenvector scalar product solutions also form a basis for sub-optimal control method for systems in which the control matrix is non-singular. This method has several advantageous features. First, in the absence of disturbances, the control vector, once obtained, remains constant until the system reaches the origin. Second, the

transition time from any point to the origin as well as the trajectory to the origin can be readily obtained. The sub-optimal isochrones can also be easily found. Third, in designing a closed-loop system using this method, the controller-computer must only solve algebraic equations and hence the control can be computed continuously. This should be contrasted with many of the present methods which require on-line solution of two-point boundary value problems and hence discrete control.

Although obtaining the optimal isochrones was computationally impossible, it was shown that the sub-optimal isochrones are tangent to the optimal isochrones at 2n points by the use of the eigenvector scalar product isochrones. Hence it appeared reasonable to conclude that the performance of the sub-optimal system was an acceptable compromise between system complexity and speed of response.

There are three major contributions of this work. First, the discovery of the eigenvector scalar product solution, second, the bounding of the optimal isochrones, and third, the design of sub-optimal control systems by the use of the eigenvector scalar product solutions.

It should be pointed out again that the development of this sub-optimal method is still incomplete. Since only systems in which the control matrix is non-singular can be treated at present, the number of practical systems to which the procedure can be applied is limited. Hence, further research is needed to extend this method to the case of singular control matrix.

7.2 Future Research Topics

As pointed out previously, the number of practical systems in which the control matrix is non-singular is very limited. Hence, in

order to make the material presented here of practical importance, it is necessary to remove this restriction. Current research is pointed toward this problem. It was noted in Chapter 5 that the eigenvector scalar product solutions do not require that $\underline{B}$ be non-singular. Therefore, these solutions can be used in creating a sub-optimal control method for systems in which the control matrix is singular. Preliminary results with second- and third-order systems indicate that this approach should be successful.

There are other extensions to the sub-optimal method which need to be made. First, it is hoped that the requirement for real eigenvalues can be removed. Second, it would be of value to extend the above method to some nonlinear problems. The most encouraging area at present is bilinear systems, in which the state and control variables are separately linear but jointly nonlinear. Because of their close relation to linear systems, it appears quite possible that the above method can be successfully applied to bilinear systems.

It is hoped that by completing these extensions that the practical significance of the method presented here will be greatly increased.

Appendix A

SUB-OPTIMAL CONTROL METHODS USING THE SECOND METHOD

A.1 Introduction

In this appendix, several methods of designing sub-optimal control systems by the use of the Second Method of Liapunov are presented. The methods presented here are not intended to be an exhaustive compilation of such methods but rather were chosen because of their relation to the material in Chapter 4.

Each of the following three sections begins with a brief discussion of the concepts or ideas underlying that method. This is followed by a short presentation of the method which is then illustrated by a numerical example. The sections conclude with a discussion of the advantages and disadvantages of each method. For each of the methods presented, the uncontrolled system is assumed to be at least stable in the sense of Liapunov.

Unfortunately, all of these methods have three basic problems: (1) they are approximate, (2) either no estimate of the approximation error is possible, or the estimate is overly conservative, and (3) it is necessary to choose a $V(x)$ for which no general procedure is presented. Hence these methods were never widely accepted.

A.2 Estimation of Transient Behavior

One of the first uses of the Second Method as a design tool was in the estimation of transient behavior[11, 12]. In particular, it was used to obtain an approximation of the settling time. By making this approximation of the settling time as small as possible, it was argued that the speed of response would be decreased. Johnson[13] has recently employed such an approach for the design of a class of sub-optimal control systems.

Consider a positive definite scalar function, $V(x)$, whose total time derivative, $\dot{V}(x)$, is negative definite. Then by the use of the Second Method, one may conclude asymptotic stability of the origin. However, although one knows that the motion tends toward the origin the rate at the origin is approached unknown. Now define η as

$$\eta = \min_{x} \frac{-\dot{V}(x)}{V(x)} \tag{A.1}$$

Then

$$\dot{V}(x) \leq -\eta V(x) \tag{A.2}$$

which may be solved to give

$$V(x(t)) \leq V(x(0))e^{-\eta t} \tag{A.3}$$

Thus, given the value of $V(x)$ at $t = 0$, an upper bound on the value of $V(x(t))$ at any time $t > 0$ can be obtained by the use of equation (A.3). Therefore from the initial state x^o the state of the system must be found within or on the surface $V(x) = V(x^o)e^{-\eta t}1$ after t_1 seconds. For an illustration of how this procedure can be used to estimate settling time, consider the following example.

Example A.1 The equations of motion for the system are

$$\begin{bmatrix} \dot{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{A.4}$$

It is desired to find an upper bound on the time that it takes the system to get from the initial condition $x^o = (1,0)$ to within the area defined by $(x_1)^2 + (x_2)^2 \leq 0.01$.

In this case it is necessary to find the largest value of K such that the surface $V(x) = K$ lies entirely within or at most tangent to the surface $(x_1)^2 + (x_2)^2 = 0.01$. See Figure A.1. Then by the use of equation (A.3), the settling time, t_s, is

$$t_s = \frac{-1}{\eta} \ln \left(\frac{K}{V(x^o)}\right) \tag{A.5}$$

However before this can be done it is necessary to find η.

Let $\dot{V}(x)$ be defined by the quadratic form $\dot{V}(x) = -x'\underline{Q}x$ where $\underline{Q}$ is a symmetric positive definite matrix. Then $V(x)$ is the quadratic from $V(x) = x'\underline{P}x$ where $\underline{P}$ is a positive definite symmetric matrix which is the unique solution of the matrix equation

$$\underline{A}'\underline{P} + \underline{P}\underline{A} = -\underline{Q} \tag{A.6}$$

Kalman and Bertram[12] have shown that η is given by

$$\eta = \text{minimum eigenvalue of } \underline{Q}\underline{P}^{-1}$$

Now let $\underline{Q}$ be

$$\underline{Q} = \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix}$$

Then by the use of equation (A.6), $\underline{P}$ is given by

$$\underline{P} = \begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix}$$

and η is equal 0.775.

For this $V(x)$, K is found to be 7.64×10^{-3}. The settling time as given by equation (A.5) is

$$t_s = \frac{-1}{0.775} \ln \left(\frac{7.64 \times 10^{-3}}{5}\right)$$
$$= 8.35 \text{ seconds}$$

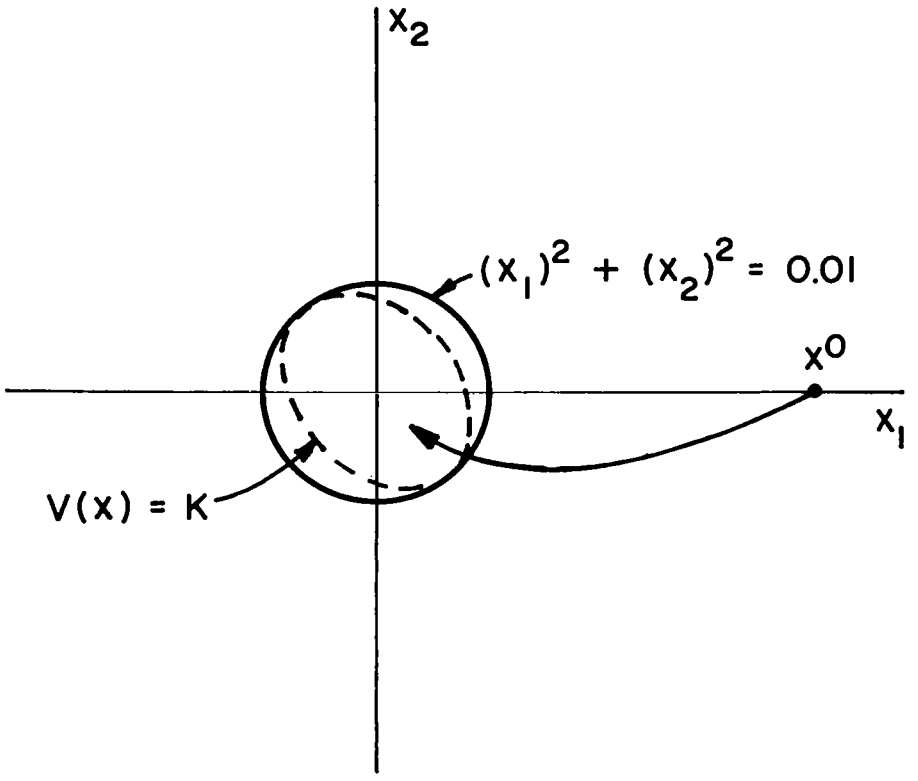


FIGURE A.1 ESTIMATION OF SETTLING TIME

This method of estimating the transient behavior of systems has several disadvantages. First, the method is approximate and no knowledge of the quality of the approximation is known. Second, the value of η and hence t_s depend on the particular V(x) used. No method of picking V(x) is known. Third, it is necessary that $\dot{V}(x)$ be negative definite. This is very difficult to attain in practice except for linear systems.

A.3 Kalman-Bertram Method

In 1960 Kalman and Bertram[12] presented a method for designing approximately time-optimal control systems. Their method was based on the knowledge that for a closed, bounded control region, U, the control vector is always on the boundary. They suggested minimizing the time derivative of V(x) arguing that this would make V(x) approach zero most rapidly, and the state of the system should reach the origin in minimum time.

Consider the system

$$\dot{x} = \underline{A}x + \underline{B}u \tag{A.7}$$

where the control region U is defined by the set of all control vectors u such that $|u_i| \leq M_i$, $i = 1,2, \ldots,n$ and M_i are positive constants. Choose arbitrarily a positive semidefinite matrix, $\underline{Q}$, and then find the positive definite matrix, $\underline{P}$, which is the unique solution of the matrix equation

$$\underline{A}'\underline{P} + \underline{P}\underline{A} = -Q \tag{A.8}$$

Now let V(x) be defined by $V(x) = x'\underline{P}x$ and $\dot{V}(x,u)$ is

$$\dot{V}(x,u) = -x'\underline{Q}x + 2u'\underline{B}'\underline{P}x \tag{A.9}$$

In order to minimize $\dot{V}(x,u)$ with respect to all admissible controls, it is necessary to minimize the second term in $\dot{V}(x,u)$. To minimize this term, each component of u must have its maximum magnitude in the direction opposite that of the corresponding component of $\underline{B}'\underline{P}x$. Therefore

$$u_i = -M_i \text{sgn}\left[(\underline{B}'\underline{P}x)_i\right] \tag{A.10}$$

As an illustration of this procedure consider the following example.

Example A.2 The equations of motion of the system are

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ -2 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

It is desired to drive this system to the origin from any initial state in minimum time.

The first step in the procedure is an arbitrary choice of $\underline{Q}$. In this case let $\underline{Q}$ be

$$\underline{Q} = \begin{bmatrix}0 & 0\\ 0 & 4\end{bmatrix}$$

in which case $\underline{P}$ as obtained from equation (A.8) becomes

$$\underline{P} = \begin{bmatrix}2 & 0\\ 0 & 1\end{bmatrix}$$

Then by the use of equation (A.10) the control vector components are found to be

$$u_1 = -M_1 \text{sgn}(2x_1)$$
$$u_2 = -M_2 \text{sgn}(2x_1 + x_2)$$

This method has several advantages. First, it provides a closed-loop solution of very simple form. The method is relatively easy to apply to high-order and multiple input systems. The control matrix $\underline{B}$ is not required to be non-singular.

The main disadvantage of the method is the lack of a procedure for choosing the $\underline{Q}$ matrix and hence $V(x)$. Since the solution depends strongly on $V(x)$, it would be highly desirable to have a procedure for choosing the "best" $V(x)$ or at least an iterative method for improving on an initial choice. Again the method provides only approximately optimum performance and no procedure for evaluating the quality of the approximation is presented. The resulting sub-optimal control system normally experiences chattering near the origin which degrades its performance.

A.4 The Nahi Method

Nahi[14] has recently presented a method of designing sub-optimal control systems based on the concept of forcing

$$\min_{u \in U} \dot{V}(x,u) \leq -K_1 V(x) - 2K_2 \sqrt{V(x)} \qquad (A.11)$$

This method was based on two arguments. First, minimizing $\dot{V}(x,u)$ would minimizing the response time. Second, forcing minimum $\dot{V}(x,u)$ to be less than or equal to $-K_1 V(x) - 2K_2 \sqrt{V(x)}$ would make the response time finite as is shown below.

The systems to be considered must be represented in the following form

$$\dot{x} = \underline{A}x + \underline{B}u \qquad (A.12)$$

where $\underline{B}$ is a non-singular matrix and the control region U is defined by

the set of all control vectors u such that $\|u\| \leq 1$. Choose arbitrarily a positive definite matrix $\underline{Q}$, and find the positive definite matrix, $\underline{P}$, which is the unique solution of the matrix equation

$$\underline{A}'\underline{P} + \underline{P}\underline{A} = -\underline{Q} \tag{A.13}$$

Now let V(x) be defined by $V(x) = x'\underline{P}x$ and then $\dot{V}(x,u)$ is

$$\dot{V}(x,u) = -x'\underline{Q}x + 2u'\underline{B}'\underline{P}x \tag{A.14}$$

In order to minimize $\dot{V}(x,u)$ with respect to all admissible controls, u must be given by

$$u = \frac{-\underline{B}'\underline{P}x}{\|\underline{B}'\underline{P}x\|} \tag{A.15}$$

Then substituting equation (A.15) for u into equation (A.14) gives

$$\min_{u \in U} \dot{V}(x,u) = -x'\underline{Q}x - 2\sqrt{x'\underline{P}\underline{B}'\underline{B}\underline{P}x} \tag{A.16}$$

Nahi[14] has shown that there exists two positive constants, K_1 and K_2 defined by

$$K_1 = \text{minimum eigenvalue of } \underline{Q}\underline{P}^{-1} \tag{A.17}$$

$$(K_2)^2 = \text{minimum eigenvalue of } \underline{P}\underline{B}\underline{B}' \tag{A.18}$$

such that the following conditions are satisfied

$$1.\quad x'\underline{Q}x \geq K_1 x'\underline{P}x \tag{A.19}$$

$$2.\quad x'\underline{P}\underline{B}\underline{B}'\underline{P}x \geq (K_2)^2 x'\underline{P}x \tag{A.20}$$

Then substituting equations (A.19) and (A.20) into equation (A.16) gives

$$\begin{aligned}\min_{u \in U} \dot{V}(x,u) &\leq -K_1 x'\underline{P}x - 2K_2\sqrt{x'\underline{P}x} \\ &\leq -K_1 V(x) - 2K_2\sqrt{V(x)}\end{aligned} \tag{A.21}$$

Now for some given initial state x(0), equation (A.21) can be solved to obtain

$$\frac{2}{K_1} \ln \left(\frac{K_1\sqrt{V(x(t))}}{2K_2} + 1\right) - \frac{2}{K_1} \ln \left(\frac{K_1\sqrt{V(x(0))}}{2K_2} + 1\right) \leq -t$$

If V(x(t)) is set equal to zero, then t becomes the transition time from x(0) to the origin, t_o.

$$t_o \leq \frac{2}{K_1} \ln \left(\frac{K_1\sqrt{V(x(0))}}{2K_2} + 1\right) \tag{A.22}$$

Hence the transition time is not only known to be finite, but also an upper bound on it is obtained. As an illustration of the above procedure consider the following example.

Example A.3 The equations of motion of the system are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & -1/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{A.23}$$

It is desired to design a sub-optimal control system which transfers the system from any initial state to the origin in a finite time. An upper bound on the transition time should also be obtained.

The first step in the procedure presented above is to arbitrarily choose a $\underline{Q}$ matrix. For this problem let $\underline{Q}$ be

$$\underline{Q} = \begin{bmatrix} 4 & 5 \\ 5 & 10 \end{bmatrix}$$

in which case $\underline{P}$ as obtained from equation (A.13) is

$$\underline{P} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Then the desired sub-optimal control as given by equation (A.15) is

$$u = \frac{1}{\sqrt{4(x_1)^2 + 4x_1x_2 + 13(x_2)^2/4}} (2x_1 + x_2, 3x_2/2)$$

From equations (A.17) and (A.18) the constants K_1 and K_2 are found to be

$$K_1 = 1.0$$

$$K_2 = 1.224$$

Then by the use of equation (A.22), the upper on the transition time is

$$t_o \leq 2 \ln \left(\frac{\sqrt{2(x_1)^2 + 2x_1x_2 + 2(x_2)^2}}{2.45} + 1 \right)$$

This method has two serious disadvantages. First, the control matrix, $\underline{B}$, must be non-singular. This, in general, is not true in practice. If $\underline{B}$ is singular, then K_2 is zero, and the transition time is infinite. Second, as pointed out in the previous section, there is no procedure for choosing the "best" $\underline{Q}$ matrix.

On the other hand, the method does provide a relatively simple closed-loop solution. The transition time is finite and an upper bound on it is readily obtained. However, there is no means of judging how close the transition time of the sub-optimal system is to the optimum.

References

1. Horowitz, Isaac, M., Synthesis of Feedback Systems, Academic Press, New York, 1963.

2. Truxal, J. G., Automatic Feedback Control System Synthesis, McGraw-Hill Book Company, Inc., New York, 1955.

3. Kalman, R. E., "The Theory of Optimal Control and the Calculus of Variations," RIAS Technical Report No. 61-3, 1961.

4. Boltyanskii, V. G., R. V. Gamkreliche, and L. S. Pontryagin, "The Theory of Optimal Processes. I. The Maximum Principle," Izv. Akad. Nauk. SSSR, Ser. Mat., Vol. 24, 1960, pp. 3-42.

5. Pontryagin, L. S., et.al., The Mathematical Theory of Optimal Processes, Interscience, New York, 1962.

6. Hahn, W., Theory and Application of Liapunov's Direct Method, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1963.

7. LaSalle, J. P., "Some Extensions of Liapunov's Second Method," IRE Trans. on Circuit Theory, December 1960.

8. Schultz, D. G., The Variable Gradient Method of Generating Liapunov Functions, with Applications to Automatic Control, Ph.D. Thesis, Purdue University, 1962.

9. LaSalle, J. P. and S. Lefschetz, Stability by Liapunov's Direct Method with Applications, Academic Press, New York, 1961.

10. Malkin, I.G., Theory of Stability of Motion, AEC Translation No. 3352, Washington, D. C.

11. Letov, A. M., Stability in Nonlinear Control Systems, Princeton University Press, Princeton, N. J., 1961.

12. Kalman, R. E. and J. E. Bertram, "Control System Analysis and Design Via the 'Second Method' of Liapunov, Part I, Continuous-Time Systems," Trans. of ASME, J. of Basic Engr., June 1960, pp. 371-393.

13. Johnson, George W., Design of Optimal Control Systems with Stability Constraints via the Direct Method of Liapunov, Ph.D. Thesis, University of Connecticut, 1963.

14. Nahi, N. E., "On Design of Time Optimal Systems via the Second Method of Liapunov," Univ. of Southern Calif., School of Engr. USCEE Report No. 103, January 1964.

15. LaSalle, J. P., "Stability and Control, " J. SIAM Control, Series A. Vol. 1, No. 1, 1962.

16. Krassovskii, N. N., "On the Theory of Optimum Control," PMM Vol. 23, No. 4, 1959, pp. 899-918.

SECTION II

NONLINEAR STABILITY OF COUPLED CORE REACTORS

Chapter 1

INTRODUCTION

A critical reactor which consists of two or more independently subcritical cores is a coupled core reactor. This description could undoubtedly be applied to many types of reactors including heterogeneous critical assemblies with the individual fuel-moderator cells treated as coupled cores. Of immediate practical importance, however, is the case of adjacent power reactors constituting a critical system, and particularly of the clustering of nuclear rocket engines.

In a coupled core reactor, the coupling effect results from neutron leakage to a given core from each of the other cores. Because leakage neutrons travel between cores in a finite time, the behavior of a given core depends not only upon processes occurring at the present time, but also upon the past history of the other cores. Herein lies the uniqueness of the problem. The set of differential equations describing the kinetics of each core contains, due to the leakage of neutrons from the other cores, source terms with the argument of the dependent variable retarded or delayed in time. Systems of equations of this type are systems with delay, systems with lag, delay-differential systems, or differential-difference systems.

The static and dynamic behavior of coupled core reactors has been investigated previously (1,3,16,17). Kinetics studies have been restricted to conventional linear analyses wherein the frequency response

of the zero power transfer function of each of the cores is examined. Additionally, conclusions on stability have been drawn by applying the Routh Test (18) to the zero power transfer function of a given core. In this method, the exponential term in the characteristic equation must be replaced by a ratio of polynomials, specifically the Pade or cut-product approximants (18). The roots of the characteristic equation for a linear system can be located exactly, and stability determined. Stability criteria for characteristic equations containing exponential terms will be discussed in Chapter 3.

It would be useful to develop a method by which the stability of coupled core systems could be readily examined. During the last three decades, researchers have found that the most universal method of investigating stability is the Second Method of Liapunov (10). It is natural, therefore, that a technique based on Liapunov's theory be derived for this problem. This approach has several advantages over the previously mentioned methods for coupled core systems. The Second Method does not require that the system be linear, and it is known that the reactor kinetics equations are nonlinear for power reactors such as rocket systems with temperature induced reactivity effects. It will be seen, moreover, that the Second Method is more useful than other methods even in solving linear problems in differential-difference systems. The Second Method yields only sufficient conditions for stability, so the results are either exact or conservative. The use of approximants for the exponential term of the characteristic, however, leads to conclusions that the system is stable when it is not in many cases.

By use of the Second Method, the extensive manipulations required to investigate each core separately will be eliminated, and stability will be discussed in terms of the system as an entity. Finally, the recent use of the Second Method in conjunction with optimal control system synthesis (6,7) dictates the understanding of the Second Method as applied to differential-difference systems toward possible further study in the area of optimum control of coupled core reactor systems.

There are no fundamental changes in Liapunov's theories in applying them to systems with delay. The theory of differential-difference equations, however, differs considerably from the theory of ordinary differential equations in questions of uniqueness, existence, and asymptotic behavior. It is necessary to demonstrate the unique properties of differential-difference equations to understand the requirements for the Liapunov function in the presence of time delay. The problem of selecting a suitable Liapunov function is one of paramount importance in studying the stability of motion of ordinary differential equations. This problem is more difficult for differential-difference equations. For example, the presence of cross products of the unknown variable with and without the retarded argument raises questions as to the sign definiteness of the functions involved. Some new concepts such as Krasovskii's Liapunov functional must be introduced.

When the difficulties mentioned above are surmounted, the usefulness of the Second Method in dealing with the stability of time delay systems in general and coupled core reactor systems in particular can be shown.

This report provides the preliminary basis for solving the problem. Chapter 2 contains the derivations of the reactor kinetics equations for coupled core systems. This development includes the representation of these equations in a form amenable to analysis by use of Liapunov's Second Method. Chapter 3 is a survey of the elements of the theory of differential-difference equations, including definitions, notation, and the question of stability or asymptotic behavior with specific examples. The Second Method of Liapunov is introduced in Chapter 4 with the emphasis on the extension of the method to time delay systems.

Chapter 2

THE KINETICS EQUATIONS FOR COUPLED CORE REACTOR SYSTEMS

Introduction

In the development that follows, a logical derivation of the kinetics equations based upon a general knowledge of the processes involved is made. The model is that of a point reactor with the various characteristic parameters representing average values with respect to space. All the neutrons that participate in the processes are of thermal energy. Because each core of the system is treated separately, the spatial effects of the variables are actually considered to some degree.

Neutron Kinetics

If there is a density n(t) of thermal neutrons with a mean lifetime ℓ_o in a given core, the net rate of disappearance of these neutrons is $n(t)/\ell_o$. The effective multiplication or total number of neutrons produced in the next generation per original neutron is k. A fraction of the produced neutrons appears some time after fission occurs. The total delayed fraction is β, consisting of the sum of the β_j, the fractions attributable to M distinct groups of atoms which decay with decay constant λ_j to produce the delayed neutrons. The density of the precursor atoms is $c_j(t)$, and the delayed neutrons appear at the same rate as the precursors decay. The net production of neutrons is

$$\frac{\varrho}{\ell} n(t) - \frac{\beta}{\ell} n(t) + \sum_{j=1}^{M} \lambda_j c_j(t), \qquad (2.1)$$

where ρ is (k-1)/k, the reactivity, and ℓ is ℓ_o/k, the effective lifetime, which is assumed to be constant. The net production of the jth group of precursor atoms is

$$\frac{\beta_j}{\ell} n(t) - \lambda_j c_j(t). \tag{2.2}$$

Expression (2.1) is incomplete because neutrons may appear from other sources which are generally external to the system. In the next section the special delayed source term for this problem is discussed. A general source S(t) is defined and Eqs. (2.1) and (2.2) are equated to the rates of change of the neutron and precursor densities, respectively. The result is

$$\frac{dn(t)}{dt} = \frac{\rho(t)}{\ell} n(t) - \frac{\beta}{\ell} n(t) + \sum_{j=1}^{M} \lambda_j c_j(t) + S(t) \tag{2.3}$$

$$\frac{dc(t)_j}{dt} = \frac{\beta_j}{\ell} n(t) - \lambda_j c_j(t).$$

The reactivity is a function of time because changes are introduced by means of an external device such as a movable neutron absorber.

Eqs. (2.3) as written are linear with a time varying parameter, $\rho(t)$. It will be seen that the reactivity can also be a function of other system variables. A nonlinear system results.

The Delayed Source

A coupled core reactor consists of N cores, and neutrons leaking from the kth core influence the ith core by contributing a source of thermal neutrons. This source is proportional at a given time to the neutron density in the kth core at a time T_{ik} earlier, where T_{ik} is the

delay time for the effective exchange of neutrons between the cores to occur. If the constant of proportionality is α_{ik}, the total source in the ith core is

$$\sum_{k=1,\neq i}^{N} \alpha_{ik}\, n_k(t - T_{ik}). \tag{2.4}$$

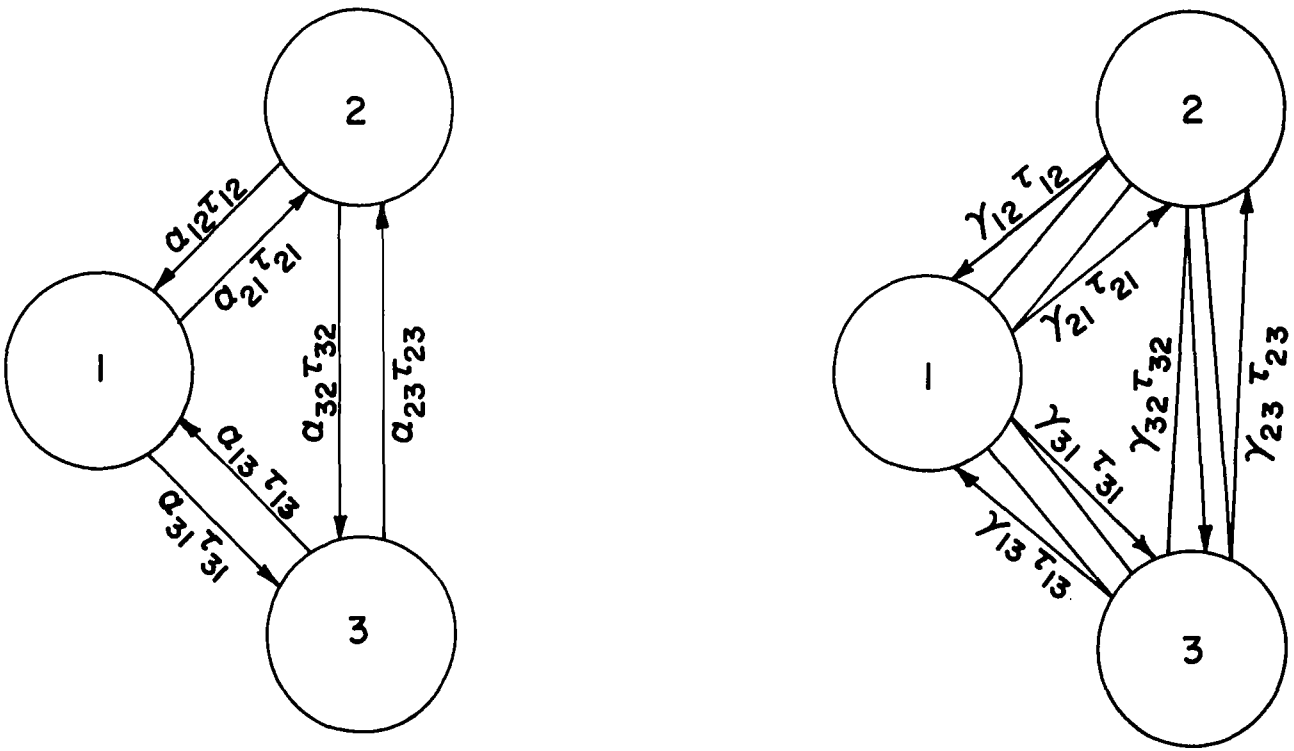
It is conceivable that there are additional terms due to reflection of leakage neutrons back to a specific core. This means that the past history of the ith core as well as of the kth core influenced the ith core. This term would be

$$\sum_{k=1,\neq i}^{N} \gamma_{ik}\, n_i(t - \tau_{ik}),$$

where γ_{ik} and τ_{ik} represent the constant of proportionality and delay time respectively for the second order effect. This process is illustrated for a three-core system in Figure 2.1. There could be even higher order effects but all except Eq. (2.4) will be neglected for now.

Power Reactors

Eqs. (2.3) represent the behavior in time of a core at zero power. In a power reactor, a coolant flows through the core removing the generated energy in the form of heat. In this case the intrinsic reactivity effects due to the temperature changes in the core appear. Normally, the change in the physical dimensions of the system, which affects neutron leakage, causes reactivity variations. In a hydrogen cooled nuclear rocket, an additional effect results from changes in the coolant density, hence in the neutron moderating properties of the system, since hydrogen is a strong moderator. In any case, the reactivity is some function of temperature.



FIRST ORDER

SECOND ORDER

FIG. 2.1 FIRST AND SECOND ORDER COUPLING EFFECTS IN A THREE-CORE SYSTEM.

A lumped parameter or point model for the heat removal process in each core will be assumed. In a rocket, especially, the parameters and temperatures of interest vary strongly throughout the core. Proper selection of mean values of variables and parameters to lead to a fairly accurate model is then of great importance.

Over a period of time, the net accumulation of energy in a core is equal to the total energy generated due to fission minus the total energy removed by the coolant. On a unit time basis, the generated energy is the power p(t), which is proportional to n(t), and the energy removed is proportional to the difference between the average core temperature and the average coolant temperature, T(t) and $T_c(t)$, respectively. The energy balance is

$$\int_0^t p(t)dt - \int_0^t H(T(t) - T_c(t))dt = MC_rT(t),$$

and differentiating with respect to time

$$MC_r \frac{dT(t)}{dt} = p(t) - H(T(t) - T_c(t)). \qquad (2.5)$$

MC_r is the product of the mass and the specific heat of the core, and H is the total heat transfer coefficient.

From a similar energy balance on the coolant with mc_c the mass heat capacity of the coolant and dm/dt or w the mass flow rate,

$$mc_c \frac{dT_c(t)}{dt} = H(T(t) - T_c(t)) - wc_cT_c(t). \qquad (2.6)$$

The reactor and coolant temperatures are proportional if $mc_c \frac{dT_c(t)}{dt}$ is negligible. This assumption is valid if the coolant is gaseous,

making mc_c small, or if the flow rate is large, making the temperature derivative insignificant. Assuming that the mass flow rate is constant, Eqs. (2.5) and (2.6) yield the heat removal equation for a given core,

$$\frac{dT(t)}{dt} = \frac{1}{MC_r} p(t) - \frac{1}{\tau} T(t), \tag{2.7}$$

where τ is the heat exchange time constant which equals $\frac{(MC_r(H + wc_c))}{Hwc_c}$. A further simplified example would result if τ and MC_r were large so that the temperature would be proportional to the power.

Some Reactivity Functions

The usual temperature dependent reactivity is, to a good approximation, a linear function of temperature. Mohler (11) shows that the effect due to hydrogen density changes in a rocket is approximately proportional to the product of the propellant flow rate and the inverse square root of the temperature.

In general

$$\rho = \rho_{oo} + \delta(t) + \rho(T) \tag{2.8}$$

where ρ_{oo} is a component of the reactivity required to maintain criticality in the steady state, $\delta(t)$ is an external reactivity input and $\rho(T)$ is the general temperature function. Usually,

$$\rho(T) = c_t T(t), \tag{2.9}$$

and for a rocket

$$\rho(T) = c_t T(t) + c_h w(T(t))^{-1/2} \tag{2.10}$$

When temperature and power are proportional,

$$\rho(T) = c_p p(t). \tag{2.11}$$

A Special State Variable Form for the Kinetics Equations

The methods of treatment and the general discussions that follow presuppose that the nth order dynamic system is representable as n first order ordinary differential equations. The kinetics equations, however, are not differential equations but differential-difference equations due to the retarded argument in the coupling source term. The limitation is, in reality, general enough to admit systems of first order differential-difference equations. This can be seen clearly from the definition of ordinary differential equations. Equations containing the derivatives of the unknowns with respect to one real variable are ordinary differential equations, therefore Eqs. (2.3) and (2.7) meet the requirements as written.

A further refinement is necessary because questions of stability will be considered with respect to some operating point. A linear change in variables such that the new variables vanish at the operating point ensures that the derivatives also vanish. This defines the equilibrium point about which stable or unstable perturbations occur.

Normalized with respect to the operating point, the transformed variables are

$$\begin{aligned} x_p(t) &= (p(t) - p_o)/p_o \\ x_c(t) &= (c(t) - c_o)/c_o \\ x_t(t) &= (T(t) - T_o)/T_o. \end{aligned} \tag{2.12}$$

The variables with the subscript o define the operating point. Substituting p(t) for n(t) with c(t) now the power due to delayed neutrons, the kinetics equations are for the ith core

$$\frac{dp_i(t)}{dt} = \frac{\rho_i}{\ell_i} p_i(t) - \frac{\beta_i}{\ell_i} p_i(t) + \sum_{j=1}^{M} \lambda_{ij} c_{ij}(t) + \sum_{k=1,\neq i}^{N} \alpha_{ik} p_k(t-T_{ik}) \tag{2.13}$$

$$\frac{dc_{ij}(t)}{dt} = \frac{\beta_{ij}}{\ell_i} p_i(t) - \lambda_{ij} c_{ij}(t)$$

$$\frac{dT_i(t)}{dt} = \frac{1}{MC_{ri}} p_i(t) - \frac{1}{\tau_i} T_i(t).$$

Setting the derivatives equal to zero, the interrelation of the equilibrium values is

$$p_{io}(\rho_{io} - \beta_i)/\ell_i + \sum_{j=1}^{M} \lambda_{ij} c_{ijo} + \sum_{k=1,\neq i}^{N} \alpha_{ik} p_{ko} = 0$$

$$\frac{\beta_{ij}}{\ell_i} p_{io} = \lambda_{ij} c_{ijo} \qquad (2.14)$$

$$\frac{1}{MC_{ri}} p_{io} = \frac{1}{\tau_i} T_{io}.$$

The solution of the first equation above for ρ_{io} and the substitution of the result into Eq. (2.8) evaluated at the operating point yields the value of ρ_{ioo} for the particular operating point chosen.

The reactivity function $\rho(T)$ separates under the transformations of Eq. (2.12) into a constant, δ_o, representing equilibrium, and a function of only the new variables, $\delta(x)$, which vanishes at equilibrium. From Eqs. (2.14) and (2.8),

$$\rho_{io} = - \sum_{k=1,\neq i}^{N} \alpha_{ik} R_k = \rho_{ioo} + \delta_{io} .$$

Therefore

$$\rho_{ioo} = - \delta_{io} - \sum_{k=1,\neq i}^{N} \alpha_{ik} R_k \qquad (2.15)$$

where the R_k are the p_{ko}/p_{io} or the "flux tilt" between the cores.

If the cores are operating at the same power initially, the R_k are unity.

The kinetics equations are transformed using Eqs. (2.12), (2.14), and (2.15). The desired forms are, with $\dot{x}(t)$ defined as $dx(t)/dt$,

$$\ddot{x}_{pi}(t) = (\delta_i(t)/\ell_i + \delta_i(x)/\ell_i)(1 + x_{pi}(t)) - (\beta/\ell)_i x_{pi}(t)$$

$$+ \sum_{j=1}^{M} (\beta_j/\ell)_i x_{cij}(t) - \sum_{k=1,\neq i}^{N} \alpha_{ik} R_k x_{pi}(t) + \sum_{k=1,\neq i}^{N} \alpha_{ik} R_k x_{pk}(t-T_{ik})$$

$$\dot{x}_{cij}(t) = \lambda_{ij} x_{pi}(t) - \lambda_{ij} x_{cij}(t) \qquad (2.16)$$

$$\dot{x}_{ti}(t) = (1/\tau_i) x_{pi}(t) - (1/\tau_i) x_{ti}(t).$$

It is seen that the derivatives vanish also for all x = -1. This result must hold because this point is the zero point for the untransformed equations.

Only the autonomous system will be considered in this study, thus $\delta_i(t)$ will be zero. Also, in the free system, the propellant mass flow rate is constant, an assumption made previously. The investigation will consist of determining how the system behaves when one or more of the variables deviates from equilibrium. This process is identical, in the linear case, with examining the roots of the transfer function, which is the ratio of the Laplace Transforms of one of the variables to that of the generalized input $\delta_i(t)$. These points will be discussed in more detail later.

If the zero point of Eqs. (2.16) represents the origin of an n-dimensional vector space, then each x represents a component of a vector which completely describes the state of the system. The variables,

in this case, are called state variables and the space is called the state space. The state variables are not necessarily physically measurable quantities. For example, the delayed neutron precursor density variable $x_c(t)$ cannot be readily measured.

The order of the system is $(N)(M + 2)$, so if there were say three cores and six groups of delayed neutrons, the order would be twenty-four. Even with no delayed neutrons and only two cores, the minimum, the order is four. The problem is formidable, if for no other reason, because of its sheer size. Experience shows that any problem greater than third order is difficult.

Eqs. (2.16) do not appear directly in vector matrix form as yet. Although this is possible, it serves no useful purpose for the general case. Some specific examples will be stated. Because one group of delayed neutrons provides a sufficiently accurate model, when delayed neutrons are considered, it will be one group. Also a two-core system is sufficient for purposes of illustration.

Power Proportional Reactivity

Using Eq. (2.11) for the reactivity,

$$\rho(T) = c_p p_o + c_p p_o x_p(t)$$

$$\delta_i(x) = \frac{a_i x_{pi}(t)}{\ell_i}$$

where a is $c_p p_o$, usually a negative quantity. The kinetics equations are, for one group of delayed neutrons,

$$\dot{x}_{pi}(t) = (a_i - b_i)x_{pi}(t) + a_i x_{pi}^2(t) + b_i x_{ci}(t) - \sum_{k=1,\neq i}^{N} c_{ik} x_{pi}(t) + \sum_{k=1,\neq i}^{N} c_{ik} x_{pk}(t - T_{ik})$$

$$\dot{x}_{ci}(t) = \lambda_i x_{pi}(t) - \lambda_i x_{ci}(t)$$

where b is β/ℓ and c_{ik} is $\alpha_{ik}R_k$.

The simplest case imaginable would result if there were no delayed neutrons and two cores. The equations are

$$\begin{aligned}\dot{x}_1(t) &= (a_1-c_{12})x_1(t) + c_{12}x_2(t-T_{12}) + a_1x_1^2(t)\\ \dot{x}_2(t) &= (a_2-c_{21})x\ (t) + c_{21}x_1(t-T_{21}) + a_2x_2^2(t)\end{aligned} \tag{2.17}$$

with x_1 and x_2 the power variables for the two cores. It would be difficult to justify this example if it were known positively that the delayed neutron source were of the same magnitude as the coupling source. With the delayed neutron effect

$$\begin{aligned}\dot{x}_{p1}(t) &= (a_1-b_1-c_{12})x_{p1}(t) + b_1x_{c1}(t) + c_{12}x_{p2}(t-T_{12})+a_1x_{p1}^2(t)\\ \dot{x}_{c1}(t) &= \lambda_1x_{p1}(t) - \lambda_1x_{c1}(t)\\ \dot{x}_{p2}(t) &= (a_2-b_2-c_{21})x_{p2}(t) + b_2x_{c2}(t) + c_{21}x_{p1}(t-T_{21})+a_2x_{p2}^2(t)\\ \dot{x}_{c2}(t) &= \lambda_2x_{p2}(t) - \lambda_2x_{c2}(t)\end{aligned} \tag{2.18}$$

If the two cores have identical nuclear properties,

$$b_1 = b_2$$

$$\lambda_1 = \lambda_2$$

$$T_{12} = T_{21}$$

but in general

$$c_{12} \neq c_{21}$$

$$a_1 \neq a_2$$

because of the flux tilting factor. The initial power, it is recalled,

is a factor in a. These quantities are equal only if the power levels are initially equal in the two cores.

These specific examples serve to demonstrate the longhand notation. The unique mathematical features of the coupled core system are seen more clearly.

Properties of the State Variable Form

The choice of variables makes linearization a comparatively easy task. The variables represent fractional deviations from equilibrium and the usual approach is to assume such deviations to be small. The higher order terms in the equations are thus neglected. For example, in Eqs. (2.17),

$$x_1^2(t) = x_2^2(t) = 0$$

A basic question of notation arises here. For Eqs. (2.17) the vector notation is

$$\underline{x}(t) = D\underline{x}(t) + C\underline{x}(t-T) + f(\underline{x}),$$

where the underlined variables are column vectors. For example,

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

D is the square matrix

$$\begin{bmatrix} (a_1-c_{12}) & 0 \\ 0 & (a_2-c_{21}) \end{bmatrix}$$

C is

$$\begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \end{bmatrix}$$

and $f(\underline{x})$ is for this example

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} x_1^2(t) \\ x_2^2(t) \end{bmatrix}.$$

These forms are possible only if the delay times between cores are equal. The general nth order case for Eqs. (2.17) is

$$\underline{x}(t) = D\ \underline{x}(t) + C\ \underline{x}(t-T)\ I\ I_c + f(\underline{x}).$$

D is again a square diagonal matrix

$$\begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ & & & & \\ 0 & & & \cdots & d_n \end{bmatrix}$$

where

$$d_i = a_i - \sum_{k=1,\neq i}^{n} c_{ik}.$$

C is the square matrix c_{ik} with c_{ik} zero if i=k, in other words, with the diagonal elements zero. I and I_c are unit square and column matrices respectively. $\underline{x}(t-T)$ is

$$\begin{bmatrix} x_1(t-T_{11}) & \cdots & x_1(t-T_{n1}) \\ & & \\ x_n(t-T_{1n}) & \cdots & x_n(t-T_{nn}) \end{bmatrix}.$$

The unit operations eliminate the diagonal elements above since they are nonexistent physically.

For the higher order problems such as Eq. (2.18) or a problem with the temperature effect, two or more equations describe each core. It is possible in this case to arrange the equations in many different ways. It is essential, however, that advantage is taken of the inherent symmetry of the system whenever possible.

Aside from notational considerations, the outstanding feature of the coupled core kinetics equations is the time delay term. The various properties of equations of this type are discussed in the following chapter.

Chapter 3

THE THEORY OF DIFFERENTIAL-DIFFERENCE EQUATIONS

Introduction and Notation

In the last section of the previous chapter, it was discovered that there could be some difficulty in arriving at a general vector notation for the equations under consideration. A completely general form for the equations is

$$\dot{\underline{x}}(t) = \underline{F}(\underline{x}(s)) \tag{3.1}$$

for all $t > t_o$, where t_o is the instant at which the solution begins. t_o will be called the initial instant, a fixed finite number.

For the ith element of Eq. (3.1), $F_i(\underline{x}(s))$ is a functional whose value depends upon the values of the function $\underline{x}(s)$, where $\underline{x}(s)$ includes the $x_i(s)$ for i=1 to n. The variable s includes all $t-T \leq s \leq t$, where T is a positive constant. Applying this notation to Eqs. (2.17),

$$F_1 = (a_1-c_{12})x_1(t) + c_{12}x_2(t-T_{12}) + a_1x_1^2(t)$$

$$F_2 = (a_2-c_{21})x_2(t) + c_{21}x_1(t-T_{21}) + a_2x_2^2(t)$$

$$\underline{F} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$$

$$T \geq T_{12}, T_{21}.$$

The vector space will be Euclidean of n dimensions, given by the symbol

E^n. For any vector $\underline{x}$ belonging to E^n, the norm is

$$\|x\| = \max |x_i|$$

where $i = 1,-----,n$. Also, if s lies on the segment $a \leq s \leq b$ in E^n, the notation

$$\|x\|^{a,b} = \sup_{a \leq s \leq b} \|\underline{x}(s)\|$$

applies for any function $\underline{x}(s)$. $\underline{x}(s)$ belongs to a class of continuous functions from the segment $a \leq s \leq b$ in a region contained in E^n.

The function $\underline{F}(\underline{y}(s))$ is continuous in time if it is a continuous function of time for $t_o \leq t \leq \gamma$ when $\underline{y}(s)$ belongs to the required class of continuous functions with $b = \gamma$ and $a = t_o - T$. $\underline{F}(\underline{y}(s))$ is locally Lipschitz with respect to $\underline{y}$ if there exists a constant L such that

$$\|\underline{F}(\underline{y}(s)) - F(\underline{z}(s))\| \leq L \, |y-z|^{t_o-T,t}$$

and $\underline{y}(s)$ and $\underline{z}(s)$ belong to a class of continuous functions on the segment $t_o - T \leq s \leq t$.

Existence and Uniqueness of Solutions

In the previous definitions the segment upon which the class of functions lies extends to $t = t_o - T$. The reason is that if the function is a solution, $\underline{x}(s)$, it must depend upon data defined for $t_o - T \leq t \leq t_o$. This is the initial data or initial function $\phi(t)$. In dealing with ordinary differential equations, it is sufficient to define an initial value of x at t_o, and the solution to the right of t_o depends only upon this initial value. The solution of a differential-difference equation, however, depends upon initial values defined over a finite time and upon

the initial instant. There could be an infinite number of $\phi(t)$s which have the same value at t_o, each resulting in a different solution to the right of t_o. Similarly, given a $\phi(t)$, if t_o changes on an absolute time scale, the solution to the right of t_o changes.

The formal proofs of uniqueness and existence are available in the literature (4,5). An example illustrates the problem. The simple first order linear differential-difference equation

$$\dot{x}(t) = -x(t-1) \tag{3.2}$$

is considered. The initial function, with $t_o = 0$, is

$$\phi(t) = 1 \quad (-1 \leq t \leq 0).$$

The solution is, for $0 \leq t \leq 1$,

$$x(t) = 1 - t,$$

and extending this to the interval $1 \leq t \leq 2$,

$$x(t) = 1 - t + 1/2(t-1)^2.$$

The solution is, by the theory of ordinary differential equations, unique for each $N-1 \leq t \leq N$. The stepwise integration continues and by induction, the general solution is

$$x_i(t) = 1 + \sum_{k=1}^{i} \frac{(-1)^k (t-k+1)^k}{k!}$$

where the ith interval is defined as

$$0 \leq t \leq 1; \; i = 1$$

$$1 \leq t \leq 2; \; i = 2$$

and so forth. If

$$\phi(t) = -t$$

the solution is

$$x(t) = -t + 1/2t^2;\ 0 \leq t \leq 1$$

and

$$x_i(t) = -t + \sum_{k=1}^{i} \frac{(-1)^{k+1}\,(t-k+1)^{k+1}}{(k+1)!}$$

Figure 3.1 displays these two solutions. From the illustrations, it is seen that choosing the second initial function corresponds exactly to moving t_o to the right by one interval. The solution to the right of t_o changes in either case.

Stability of Differential-Difference Equations

Because the system is autonomous, only asymptotic stability will be considered. Physically, the system is asymptotically stable if when perturbed from the equilibrium state, it returns to the equilibrium state. The stability property for equations with time delay is stated in the following definition.

Definition 3.1 Asymptotic Stability

The origin of Eq. (3.1) is stable if for every $\epsilon > 0$ there exists a $\delta > 0$ such that when

$$\|\phi\|^{t_o-T,t_o} \leq \delta$$

the inequality

$$\|\underline{x}(t,t_o,\phi)\| < \epsilon$$

holds for all $t > t_o$. If in addition to these conditions,

$$\lim_{t \to \infty} \|\underline{x}(t,t_o,\phi)\| = 0$$

and

$$\|\underline{x}(t,t_o,\phi)\| < H_1$$

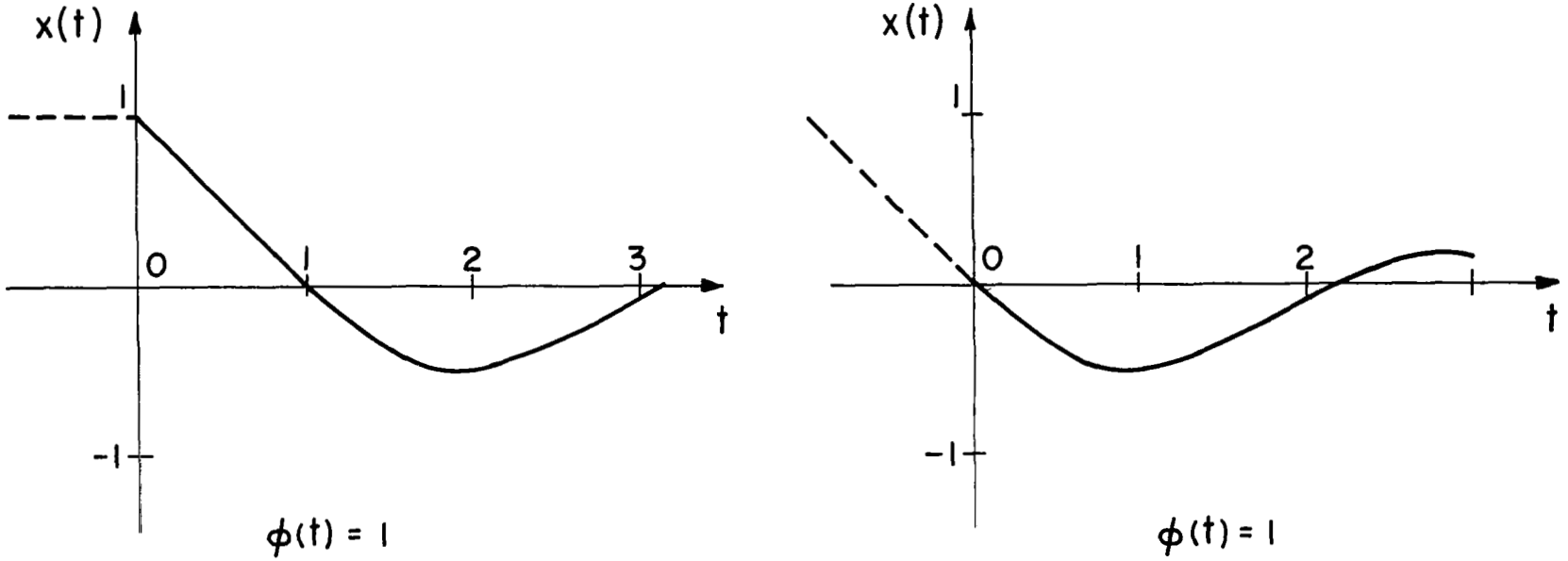


FIG. 3.1 SOLUTION OF $\dot{x}(t) = -x(t-1)$

for all initial functions

$$\|\phi\|^{t_o-T,t_o} \leq H_o,$$

then the origin or null solution ($x = 0$) is asymptotically stable.

This means that all solutions starting in a region of E^n, defined by H_o, remain in a region H_1 and approach the origin as time goes to infinity. Definition 3.1 is therefore not unlike the definition of asymptotic stability for ordinary differential equations. The difference is that the initial functions instead of initial points confine themselves to the region H_o.

Definition 3.1 may actually be too general in that a large class of initial functions is admitted. Consideration of only a rather restricted class of functions might lead to a broader applicability of stability conditions in practical problems. However, Krasovskii (8) shows rigorously that a satisfactory restriction is that $\phi(t)$ satisfies a Lipschitz condition. This allows a sufficiently large number of $\phi(t)$ especially in the practical case under consideration where an infinite discontinuity in the initial function would not be expected.

The Zeros of the Characteristic Equation

If Eq. (3.1) is linear and rewritten as

$$H(p) = \mathcal{L}\,(\dot{\underline{x}}(t) - \underline{F}(\underline{x}(s))$$

where $\mathcal{L}$ denotes the Laplace transform. $H(p)$ is the characteristic equation. With zero initial conditions on the $x_1(t)$, the roots of $H(p)$ are the poles of the Laplace transform solution of $x(t)$. Positive roots

indicate exponentially increasing solutions and negative roots indicate exponentially decreasing solutions. Therefore, a condition for asymptotic stability is that all the zeros of H(p) are negative in the real part. In general, $H(p) = H(p, e^{p})$ since $\mathcal{L}\ (x(t-T) = X(p)e^{-pT}$. These functions are called exponential polynomials. The stability criteria for exponential polynomials are stated by Bellman (2).

H(p) is multiplied by a sufficient power of e^{pT} to eliminate all negative exponential terms. If the product of the highest order of p and of e^{pT} does not appear, there are an infinite number of roots with arbitrarily large real parts. This product is called the principal term whose absence ensures instability. This is seen in the following example.

$$H(p) = e^{p} - p$$

and

$$p = \sigma + j\omega.$$

Solving for H(p) = 0,

$$\cos\omega = \sigma e^{-\sigma}$$

$$e^{\sigma} = \omega/\sin\omega.$$

If σ is arbitrarily large, $\cos\omega$ approaches zero, or $\omega = 2n\pi + 1/2\pi$. Then from the second equation, since $\sin\omega$ approaches unity,

$$\sigma = \log_{e}(2n\pi + 1/2\pi).$$

Thus σ is arbitrarily large and positive as n increases. The stability criteria are stated in Theorem 3.1.

Theorem 3.1 Stability Criteria for Linear Differential-Difference Equations from the Characteristic Equation

H(p) is an exponential polynomial with a principal term. $p = j\omega$ and $H(j\omega)$ is separable into $F(\omega) + jG(\omega)$. In order that

the zeros of $H(j\omega)$ lie to the left of the imaginary axis (have negative real parts) condition (a) or (b) must be satisfied:

(a) The zeros of $F(\omega)$ are real and for each zero, ω_o,

$$\frac{dF(\omega_o)}{d\omega} \cdot G(\omega) < 0.$$

(b) The zeros of $G(\omega)$ are real and for each zero, ω_o,

$$\frac{dG(\omega_o)}{d\omega} \cdot F(\omega_o) > 0.$$

The first order equation

$$\dot{x}(t) = -ax(t) - bx(t-T) \tag{3.3}$$

will be examined via the Second Method of Liapunov, so it is convenient to use this example and obtain the exact result. The characteristic in the proper form is

$$pe^{pT} + e^{pT} + b = 0.$$

The presence of the principal term is noted. Condition (b) of Theorem 3.1 is applied.

$$F = -\theta\sin\theta + aT\cos\theta + bT$$

$$G = \theta\cos\theta + aT\sin\theta$$

$$dG/d\omega = -\theta\sin\theta + \cos\theta + aT\cos\theta$$

where $\theta = \omega T$. $\theta = 0$ is a root of G, so

$$aT > -bT.$$

For all other roots, the parametric equations are,

$$aT = -\theta\cot\theta$$

$$bT < + (\theta^2 + (aT)^2)^{1/2}.$$

An approximate answer results if the lowest order Pade approximant (18) is used.

$$e^{-pT} \approx (2-pT)/(2+pT).$$

The Routh Test is applied to the resulting characteristic, and the range of parameters for stability is,

$$aT > -bT$$

$$bT < 2 + aT$$

which is an overestimate. These results appear in Figure 3.2.

From the solution to Equation (3.2) it could be reasoned that there probably exists some value of the delay time for which the system becomes unstable. The series form of the solution approaches that for a simple negative exponential solution for a sufficiently small value of the delay. The exact value can be found from Theorem 3.1.

The stability question will be pursued further in terms of Liapunov's Second Method in the next chapter.

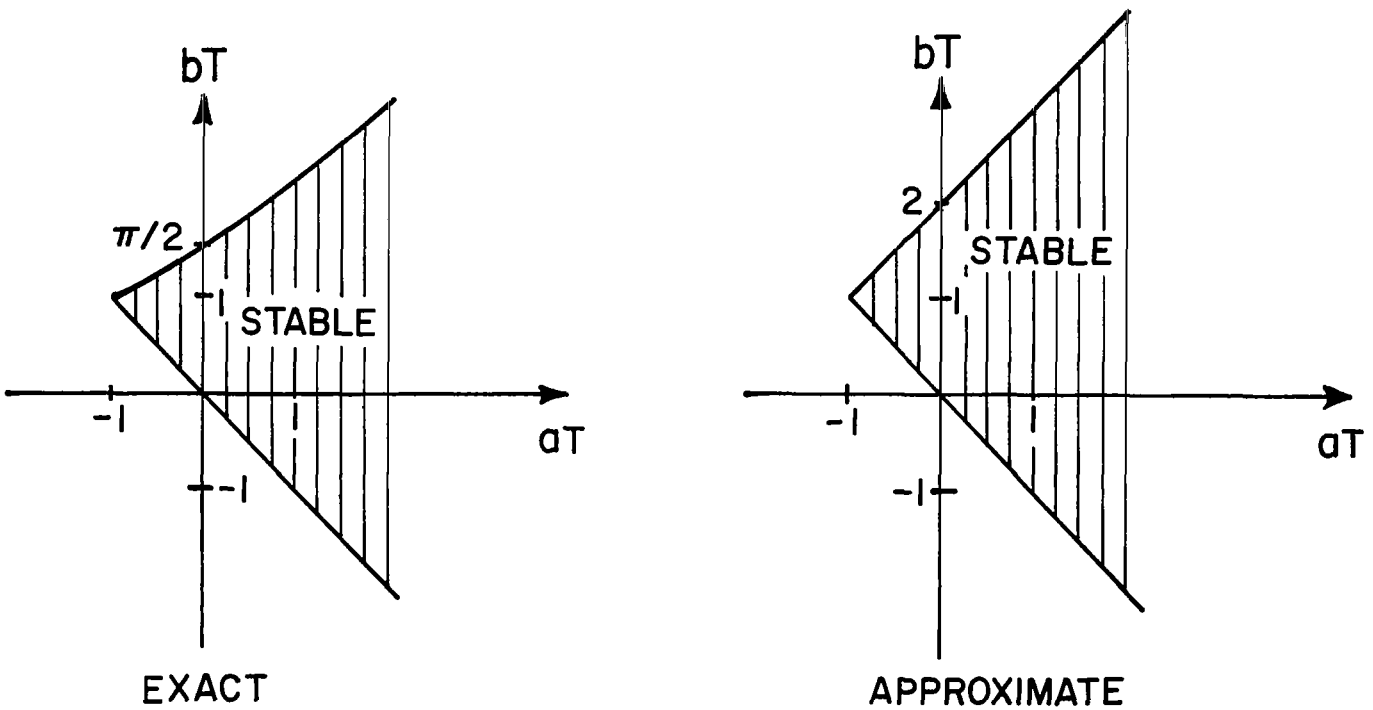


FIG. 3.2 PARAMETERS FOR ASYMPTOTIC STABILITY OF

$$\dot{x}(t) = -ax(t) - bx(t-T)$$

Chapter 4

LIAPUNOV'S SECOND METHOD FOR SYSTEMS WITH TIME DELAY

Review of the Second Method for Ordinary Systems

The idea of the Second or Direct Method of Liapunov is to determine stability for a system without a knowledge of the solutions of the system. The tool for accomplishing this is the Liapunov function $v(\underline{x})$, a scalar function of the vector $\underline{x}(t)$ for the general autonomous system

$$\dot{\underline{x}}(t) = A(\underline{x})\underline{x}(t). \tag{4.1}$$

$v(\underline{x})$ is a positive definite function, that is, it has the properties

(a) $v(\underline{x})$ is continuous with continuous first partials in a region H about the origin of E^n,

(b) $v(0) = 0$,

(c) $v(\underline{x})$ is positive in H except at the origin.

Also, $v(\underline{x}) \rightarrow \infty$ as $\|x\| \rightarrow \infty$. This ensures that $v(\underline{x}) =$ a constant represents a series of closed surfaces about the origin. The state of the system lies on successively smaller $v(\underline{x})$ toward the origin if the system is asymptotically stable. Accordingly, the basic stability theorem is:

Theorem 4.1 Asymptotic Stability

If there exists in some region H about the origin of E^n a Liapunov function $v(\underline{x})$, and if $\dot{v}(\underline{x})$ is negative definite in H, the origin is asymptotically stable.

The region H could be arbitrarily large in which case the system is globally assymptotically stable. In many nonlinear problems, however,

H is finite.

Selecting a suitable $v(\underline{x})$ for a given problem is of importance, and several methods are available for accomplishing this task. A basic $v(\underline{x})$ for Eq. (4.1) is the quadratic form

$$v(\underline{x}) = \sum_{i,j=1}^{n} a_{ij}x_i(t)x_j(t). \tag{4.2}$$

$\dot{v}(\underline{x})$ must be negative definite, a condition reached by proper selection of the a_{ij}s. Sylvester's Theorem gives the conditions for the positive-definiteness of Eq. (4.2). This will prove useful later.

Theorem 4.2 Sign Definiteness of a Quadratic Form

The function $\sum_{i,j=1}^{n} a_{ij}x_ix_j$ is positive definite if the successive principal minors of the symmetric determinant $|a_{ij}|$ are positive.

A second order example is

$$v(\underline{x}) = a_{11}x_1^2(t) + 2a_{12}x_1(t)x_2(t) + a_{22}x_2^2(t),$$

which by Theorem 4.2 is positive definite if

$$a_{11} > 0$$

$$a_{11}a_{22} - a_{12}^2 > 0.$$

A useful approach for obtaining the a_{ij} for the linear case is to constrain $\dot{v}(\underline{x})$ along solutions of the system to be

$$-\sum_{i=1}^{n} x_i^2(t).$$

For a nonlinear problem, this approach is used for the linearized equations, then $\dot{v}(\underline{x})$ is found along solutions of Eq. (4.1). This leads

to an estimate of the size of the region H in which the system is asymptotically stable.

Extension of the Second Method to Time Delay Systems

To illustrate the difficulties involved in this problem, an attempt is made to extend the Second Method directly to the first order system (3.3). $v(\underline{x})$ is chosen to be

$$v(\underline{x}) = x^2(t) \tag{4.3}$$

so

$$\dot{v}(\underline{x}) = 2x(t)\dot{x}(t) = -2ax^2(t) - 2bx(t)x(t-T). \tag{4.4}$$

The method fails. No conclusion can be drawn as to the sign definiteness of Eq. (4.4). An examination of Figure 3.2, however, reveals that the system is indeed asymptotically stable for a known range of parameters.

An idea for a new method comes from the functional system representation

$$\dot{\underline{x}}(t) = \underline{F}(\underline{x}(s)) \tag{4.5}$$

A natural approach is to seek a functional $V(\underline{x}(s))$ of the vector $\underline{x}(s)$ for $t-T \leq s \leq t$ rather than a function $v(\underline{x}(t))$ of the vector $\underline{x}(t)$. The functional will be called $V(\underline{x})$.

The Liapunov Functional

Several authors examine the stability of Eq. (3.3) by use of the functional

$$V(\underline{x}) = x^2(t) + a\int_{t-T}^{t} x^2(s)\,ds.$$

A more general form is

$$V(\underline{x}) = x^2(t) + \mu\int_{t-T}^{t} x^2(s)\,ds \tag{4.6}$$

where μ is a constant. This particular functional form is due to Krasovskii. Apparently some experience is required in choosing a proper functional just as it is in choosing a Liapunov function. The basic Liapunov approach is unchanged, so if $V(\underline{x})$ is positive definite and $\dot{V}(\underline{x})$ is negative definite along solutions of Eq. (4.5), the system is stable. $V(\underline{x})$ is positive definite, in fact,

$$x^2(t) \leq V(\underline{x}) \leq (1 + \mu T) \left\| \underline{x}(s) \right\|^{2^{t-T,t}} \tag{4.7}$$

if $\mu > 0$, where $\mu T \left\| \underline{x}(s) \right\|^{2^{t-T,t}}$ is the largest value the integral assumes in s. Differentiating Eq. (4.6),

$$\dot{V}(\underline{x}) = 2x(t)\dot{x}(t) + \mu x^2(t) - \mu x^2(t-T),$$

and from Eq. (3.3),

$$\dot{V}(\underline{x}) = -(2a - \mu)x^2(t) - 2bx(t)x(t-T) - \mu x^2(t-T). \tag{4.8}$$

Eq. (4.8) is a quadratic form in x(t) and x(t-T), so from Sylvester's Theorem, $\dot{V}(\underline{x})$ is negative definite if

(a) $2a - \mu > 0$

(b) $\mu(2a - \mu) - b^2 > 0.$

The maximum value of (b) occurs when $\mu = a$, and since μ must be > 0, a also is > 0. The range of parameters for asymptotic stability is therefore

$$a^2 - b^2 > 0,$$

or

$$a > |b|, \quad a > 0. \tag{4.9}$$

This region is shown, along with the exact boundaries, in Figure 4.1.

From the observations above, the basic stability theorem is modified as follows.

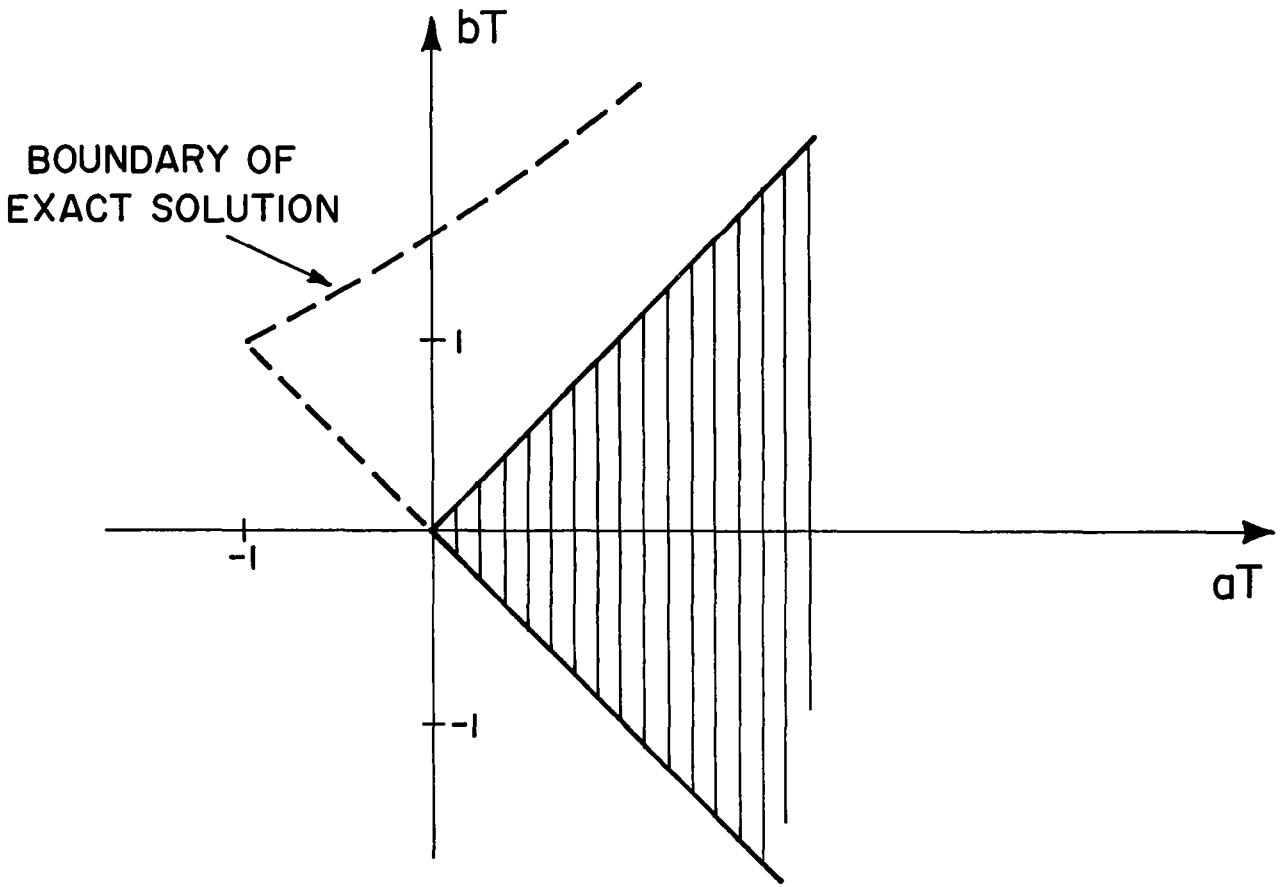


FIG. 4.1 REGION OF STABLE PARAMETERS OF $\dot{x}(t) = -ax(t) - bx(t-T)$ BY THE SECOND METHOD — $V(x(s)) = x^2(t) + a\int_{t-T}^{t} x^2(s)\,ds$

Theorem 4.3 Asymptotic Stability of a System With Delay Via a Liapunov Functional

If for the region $\underline{x}(s) < H$ and $t \geq 0$, for the system (4.5), there exists a functional $V(\underline{x}(s))$ such that

(a) $V(0) = 0$,

(b) $V(\underline{x}(s))$ is continuous in time and locally Lipschitz with respect to $\underline{x}(s)$,

(c) $V(\underline{x}(s)) \geq w(\underline{x}(t))$, where w is a positive continuous function in H,

(d) $\dot{V}(\underline{x}(s)) \leq -w_1(\underline{x}(t))$ along solutions of the system, where w_1 is a positive continuous function in H,

then the system is asymptotically stable for $t > 0$ as defined in Definition 3.1.

Condition (b) implies continuity of $V(\underline{x})$ with respect to $\underline{x}(s)$, or the existence of the derivative $\dot{V}(\underline{x})$. Condition (c) defines $V(\underline{x})$ as positive definite, and Condition (d) defines $\dot{V}(\underline{x})$ as negative definite. The functional (4.6) meets all the requirements stated in Theorem 4.3 if the conditions (4.9) are true. The function $w(\underline{x}(t))$ is $x^2(t)$. The function $w_1(\underline{x}(t))$ exists, but it is convenient to determine negative definiteness by use of Sylvester's Theorem.

From Eq. (4.7), $V(\underline{x})$ also has an upper bound. The condition

$$V(\underline{x}(s)) \leq W(\|\underline{x}(s)\|^{t-T,t}); \; W(0) = 0 \qquad (4.10)$$

could replace Condition (a) in Theorem 4.3. Inequality (4.10) is a stronger condition than (a) and the existence of the function W leads to the conclusion of uniform asymptotic stability. If a system is uniformly asymptotically stable, it is stable in the sense of Definition 3.1,

independent of t_o. Driver (4) proves the uniform property of the stability under these conditions.

The functional of Eq. (4.6) for the general nth order system would be

$$V(\underline{x}) = v(\underline{x}) + \sum_{i,j=1}^{n} \mu_{ij} \int_{t-T}^{t} x_j^2(s)ds \qquad (4.11)$$

where the subscripts i and j indicate that the ith equation may contain all the variables with arguments delayed by several times T_{ij}. The function $v(\underline{x})$ is the normal Liapunov function for the system with all the delayed terms zero.

The Liapunov Function for Time Delay Systems

The original attempt to solve the stability problem using a Liapunov function failed because $\dot{v}(x)$ was not negative definite in Eqs. (4.3) and (4.4). The solution to Eq. (3.3) is, by stepwise integration and induction,

$$x_n(t) = (-b/a)^n$$

$$+(1 + b/a) \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{(-1)^{k-1}(t-(k-1)T)^{j-1}(b)^{j-1}(b/a)^{k-j}e^{-a(t-(k-1)T)}}{(j-1)!}$$

for the segment $(n-1)T \leq t \leq nT$, and for $\phi(t) = 1$. The solution appears in Figure 4.2 for $a = b = T = 1$. The solution exhibits an oscillatory behavior which appears only for the second order system of ordinary differential equations. With time delay, $v(\underline{x})$ for the first order system is $x^2(t)$ which also oscillates, so $\dot{v}(\underline{x})$ is positive for certain times. The functional and function and their derivatives are plotted in Figure 4.3 for this example. This figure also demonstrates that the functional is the natural approach to the problem.

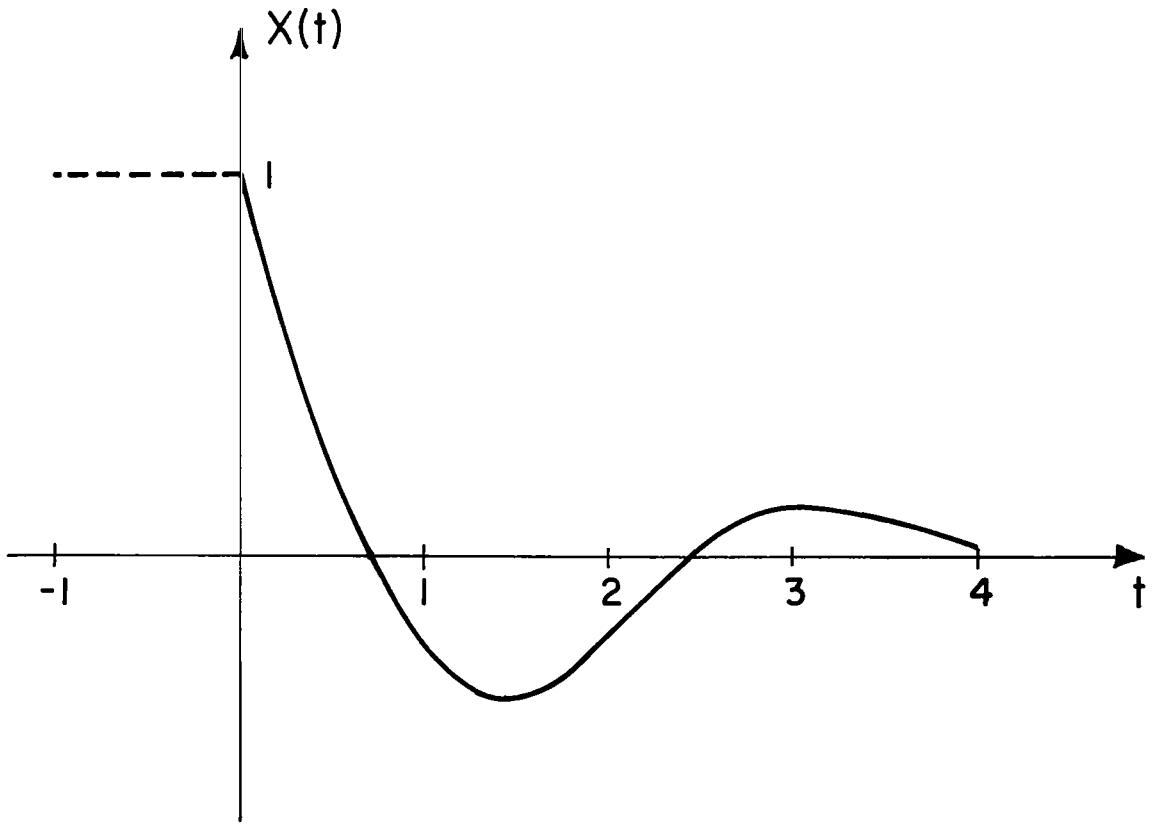
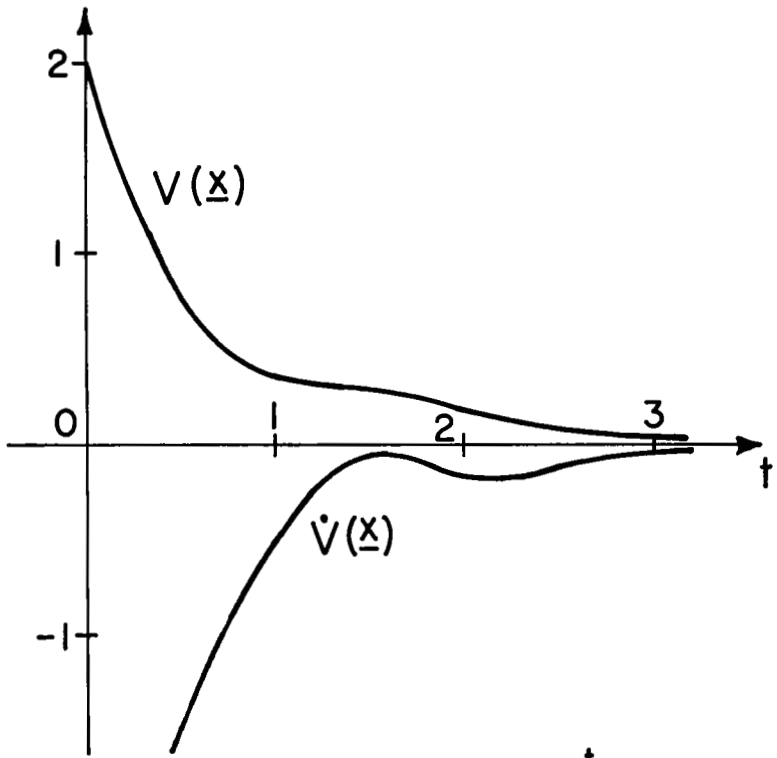
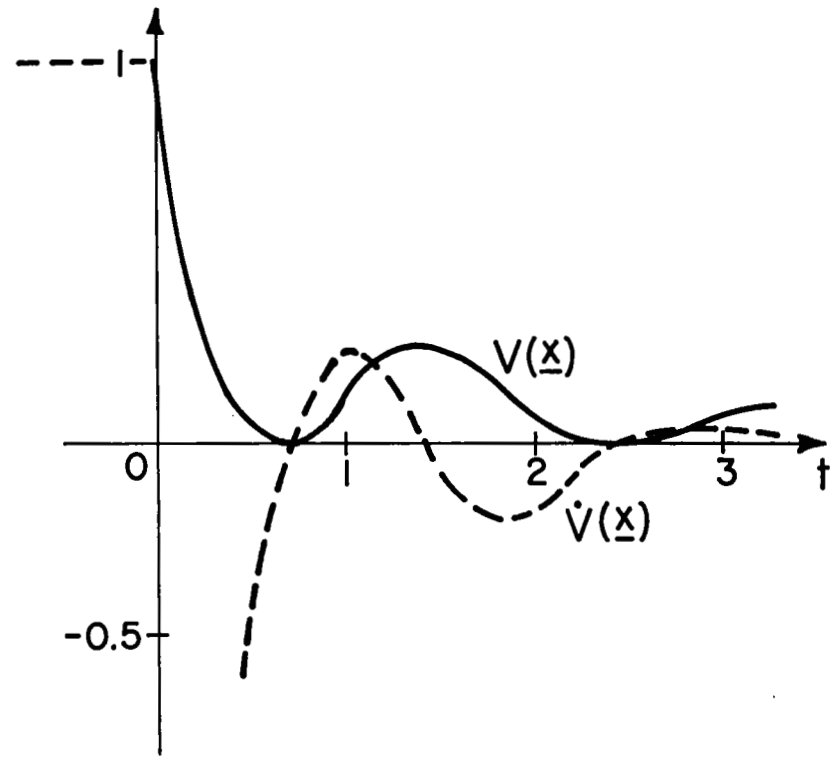


FIG. 4.2 SOLUTION OF $\dot{x}(t) = -x(t) - x(t-1)$



FUNCTIONAL = $x^2(t) + \int_{t-1}^{t} x^2(s)\,ds$



FUNCTION = $x^2(t)$

FIG. 4.3 LIAPUNOV FUNCTIONS AND FUNCTIONALS FOR

$\dot{x}(t) = -x(t) - x(t-1)$

Despite the apparent inconsistencies, the system is known to be asymptotically stable not only due to previous considerations, but also due to intuitive feelings resulting from an examination of Figure 4.2. The system appears to be returning to equilibrium. The following Lemma sets forth an additional consideration.

Lemma 4.1 Asymptotic Properties of $v(\underline{x})$

For all initial functions $\|\phi(t)\|^{t_o-T,t_o} \leq H_o$, let $\|\underline{x}(s)\| < H$ for all $t_o \leq t < \infty$. A function $v(\underline{x}(t))$ is bounded uniformly for all $\|\underline{x}(t)\| < H$, $t \geq t_o$. Suppose that $v(\underline{x}(t))$ has the property that for some $\gamma > \gamma_o$ there exists $\alpha(\gamma)$, $\theta(\gamma) > 0$ such that along solutions of the system, Sup $\dot{V}(\underline{x}) < -\alpha(\gamma)$ for $t \geq t_o + \theta(\gamma)$ for all solutions satisfying $v(\underline{x})(0)) \leq \gamma$. Then

$$\lim_{t \to \infty} \sup v(\underline{x}(\phi, t_o)) \leq \gamma_o$$

independent of $\phi(t)$.

This means that there is some $t > t_o + \theta(\gamma)$ beyond which $v(\underline{x})$ decreases monotonically. The idea is incorporated into a theorem.

Theorem 4.4 Asymptotic Stability of a System With Delay Via a Liapunov Function

If for the region $\|\underline{x}(t)\| < H$ and $t \geq -T$ for the system (4.5) there exists a function $v(\underline{x}(t))$ such that

(a) $v(\underline{x}) \leq W(\underline{x})$ where W is continuous in H and $W(0) = 0$,

(b) $v(\underline{x})$ is continuous in time and locally Lipschitz with respect to $\underline{x}$,

(c) $v(\underline{x}) \geq w(\underline{x})$ where w is continuous and positive (v is positive definite),

(d) there exists a continuous function $f(r) > r$ for all $r > 0$ and a positive continuous function $w_1(\underline{x}) > 0$ such that

$$\dot{v}(\underline{x}) \lesssim -w_1(\underline{x}) \text{ (negative definite)}$$

for $t \geq 0$, and

$$v(\underline{x}(s)) < f(v(\underline{x}(t))) \text{ for all } t-T \leq s \leq t$$

then the system is uniformly asymptotically stable for $t > 0$.

If Condition (a) is relaxed to $v(\underline{0}) = 0$, the conclusion is asymptotic stability. Condition (d) requires that the function $v(\underline{x})$ is decreasing monotonically to the right of $t_o + \theta$, in accordance with Lemma 4.1.

As an example, for Eq. (3.3), $v(\underline{x})$ is again $x^2(t)$ and

$$\dot{v}(\underline{x}) = -2ax^2(t) - 2bx(t)x(t-T).$$

$f(v)$ will be v/q where $0 < q < 1$. Then by Theorem 4.4,

$$x^2(t-T) < x^2(t)/q \tag{4.12}$$

Clearly

$$\dot{v}(\underline{x}) \leq -2(ax^2(t) - |bx(t)x(t-T)|) \tag{4.13}$$

and from Eq. (4.12)

$$|x(t)| > q^{1/2} |x(t-T)|. \tag{4.14}$$

If in inequality Eq. (4.13) $x(t-T)$ is replaced by $x(t)/q^{1/2}$, the resulting inequality

$$|b \|x(t)\| x(t-T)| < b \|x(t)\| x(t)|/q^{1/2}$$

leads to the conclusion

$$\dot{v}(\underline{x}) \leq -2(a - |b|/q^{1/2})x^2(t). \tag{4.15}$$

The right hand side of inequality (4.15) is negative definite if

$$a > 0$$

$$a > |b|/q^{1/2}$$

or

$$a \gtrsim |b|;\ a > 0$$

since $q^{1/2}$ can be arbitrarily close to unity. This is the same result as that obtained by use of the functional.

A Different Functional Form for the System Equations

If a system of ordinary differential equations is written in a different form, a different stability result is found if the same Liapunov function is used. For example, if the system is

$$\begin{aligned} \dot{x}_1(t) &= -ax_1(t) - bx_2(t) \\ \dot{x}_2(t) &= -cx_2(t) - dx_1(t) \end{aligned} \tag{4.16}$$

the equations may be rewritten in the phase variable form

$$\ddot{x}(t) + (a + c)\dot{x}(t) + (ac - bd)x(t) = 0$$

or

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -(a + c)x_2(t) - (ac - bd)x_1(t). \end{aligned} \tag{4.17}$$

If $v(\underline{x}) = x_1^2(t) + x_2^2(t)$,

$$\dot{v}(\underline{x})_{4.16} = -ax_1^2(t) - (ab + cd)x_1(t)x_2(t) - cx_2^2(t)$$

$$\dot{v}(\underline{x})_{4.17} = -(ac - bd - 1)x_1(t)x_2(t) - (d + c)x_2^2(t).$$

$\dot{v}(\underline{x})_{4.16}$ indicates stability for

$$a > 0$$

$$ac - (ab + cd)/4 > 0.$$

$\dot{v}(\underline{x})_{4.17}$ is indefinite, so no conclusion on stability results. The resulting ranges of parameters can be superimposed in this method to perhaps yield an answer which is better than either individual result.

This idea can be extended even to the first order case for differential-difference equations. The equations are rewritten using an integral representation

$$x(t-T) = x(t) - \int_{t-T}^{t} \dot{x}(s)ds. \tag{4.18}$$

Eq. (3.3), as an example, takes the form

$$\dot{x}(t) = -(a + b)x(t) - ab \int_{t-T}^{t} x(s)ds - b^2 \int_{t-T}^{t} x(s-T)ds \tag{4.19}$$

if Eq. (4.18) replaces x(t-T). If $v(\underline{x}) = x^2(t)$,

$$\dot{v}(\underline{x}) = -2(a + b)x^2(t) - 2ab \int_{t-T}^{t} x(s)x(t)ds - 2b^2 \int_{t-T}^{t} x(s-T)x(t)ds \tag{4.20}$$

$$\leq -2((a + b)x^2(t) - |ab| \int_{t-T}^{t} |x(s)\| x(t)| \, ds - b^2 \int_{t-T}^{t} |x(s-T)\| x(t)|ds)$$

Under the conditions of Theorem 4.4, the integrals in Eq. (4.20) must be less than the maximum value of the integral over $t-T \leq s \leq t$, which is

$$Tx^2(t)/q^{1/2}.$$

The constant q is different for each integral, and for expediency, each will be assumed to be unity. It is recalled that q may be arbitrarily close to unity.

Inequality (4.20) is, therefore,

$$\dot{v}(\underline{x}) \leq -2((a + b) - |ab|T - b^2T)x^2(t), \tag{4.21}$$

which consists of two cases,

$$\begin{aligned} &(a) \quad \dot{v}(\underline{x}) \leq -2(a(1 - bT) + b(1 - bT))x^2(t); \; a \;\; 0, \; b \;\; 0 \\ &(b) \quad \dot{v}(\underline{x}) \leq -2(a(1 + bT) + b(1 + bT))x^2(t); \; a \;\; 0, \; b \;\; 0. \end{aligned} \tag{4.22}$$

The right sides of the inequalities (4.22) are negative definite if

$$\begin{aligned} &(a) \quad aT > 0, \; 0 < bT < 1 \\ &(b) \quad -bT(1 - bT)/(1 + bT) < aT < 0, \; b > 0. \end{aligned} \tag{4.23}$$

The Regions (4.23) appear in Figure 4.4. This result is combined with the previous result $a > |b|$, and compared with the exact result. The answer is considerably improved and reasonably close to the exact answer. The maximum value of $-aT$ in Eq. (4.23,b) occurs at $b = \sqrt{2} - 1$ and $-aT = 2\sqrt{2} - 3$.

Conclusions

The simple first order linear example used here illustrates the mechanics of the method. The decision to be made is whether to use the functional or function approach. While the Liapunov functional is the natural tool, the function seems to lead to better results with a minimum of effort in the problem examined.

New results could be achieved through the choice of a new functional form. It is difficult to guess just what form would be useful. Krasovskii has introduced a functional involving a double integration which, for Eq. (3.3) yields a region

$$aT > 0, \quad 0 < bT < 1.$$

Due to the extra manipulations, this approach is not interesting except in cases in which all the variables except the derivative have the retarded

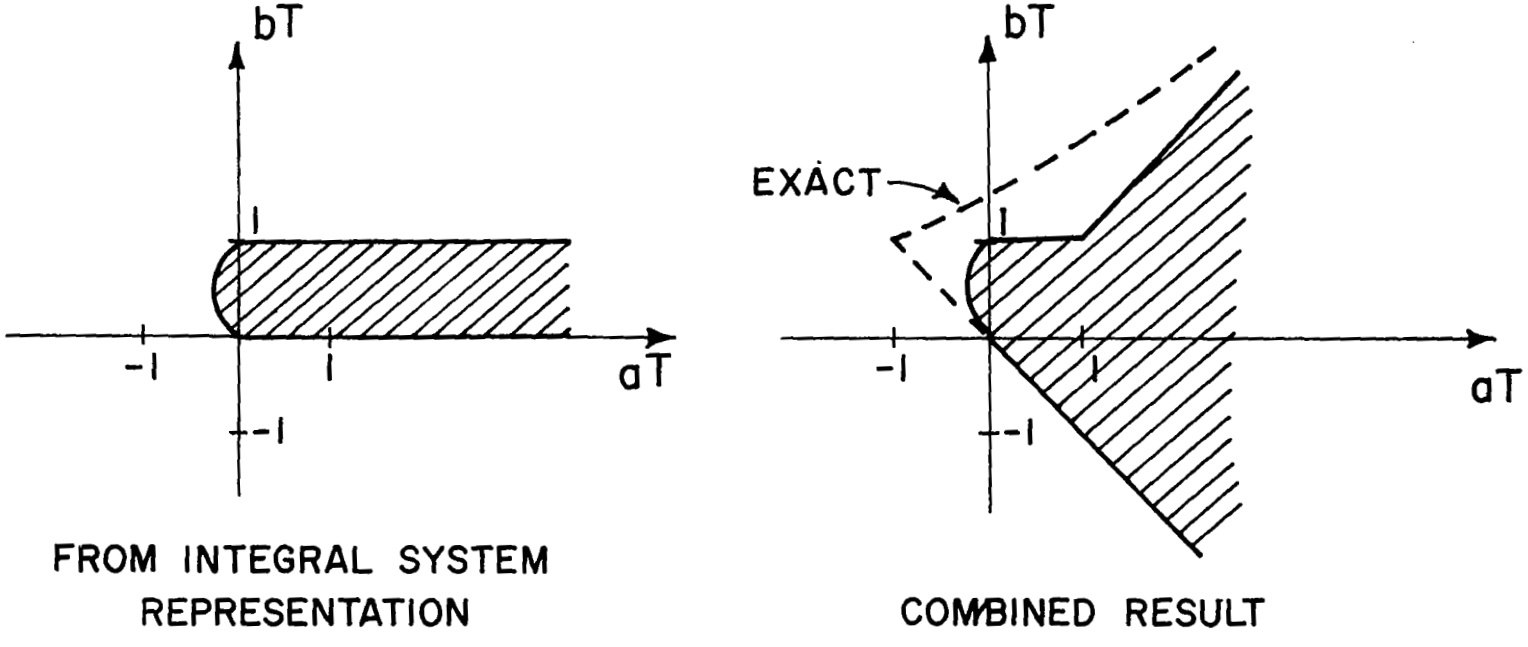


FIG. 4.4 REGIONS OF STABLE PARAMETERS OF

$$\dot{x}(t) = -ax(t) - bx(t-T)$$

argument. In this case the functional used here fails to solve the problem.

Razumikhin (13) has arrived at the regions in Figure 4.4 by a slightly different approach. Only the results are given in his paper, and his figures are grossly exaggerated and overoptimistic. In the same paper, Razumikhin gives the results of a third order linear problem. These results are somewhat sketchy, but this is the most difficult problem worked in the literature.

Chapter 5

FURTHER STUDY

The next step is to apply these methods to the coupled core reactor system equations. Some linear problems will be worked using the Second Method, and the results compared to the exact solutions. This will demonstrate the usefulness of the Second Method in solving linear time delay problems.

A method will be developed to deal with the nonlinear system. It is felt that a suitable approach would be to find Liapunov functions for the linearized system, then to estimate the region of stability by calculating $\dot{v}$ along solutions of the nonlinear system. Stability results will be given in terms of the parameters of interest (delay times, coupling coefficients, and flux tilt) and in terms of the regions of stability in the state space.

REFERENCES

1. Baldwin, G. C., "Kinetics of a Reactor Composed of Two Loosely Coupled Cores," Nuclear Science and Engineering, 6, pp. 320-327, (1959).

2. Bellman, Richard, and Cooke, K. L., Differential-Difference Equations, The RAND Corporation, Mathematics in Science and Engineering Series, 6, Academic Press, (1963).

3. Danofsky, R. A., and Uhrig, R. E., "The Kinetic Behavior of the Coupled Regions of the UTR-10 Reactor," Nuclear Science and Engineering, 16, pp. 131-133, (1963).

4. Driver, Rodney D., "Existence and Stability of Solutions of a Delay-Differential System," Archive for Rational Mechanics and Analysis, 10, pp. 401-426, (1962).

5. Hale, J. K., "Asymptotic Behavior of the Solutions of Differential-Difference Equations," The Martin Company RIAS Technical Report No. 61-10, Baltimore, Maryland, (1961).

6. Kalman, R. E., and Bertram, J. E., "Control System Analysis and Design Via the 'Second Method' of Liapunov--II--Discrete Time Systems," Transactions of the ASME, pp. 394-400, (June 1960).

7. Koepke, R. W., "On the Control of Linear Systems With Pure Time Delay," Proceedings of the 1964 Joint Automatic Control Conference, Stanford University, pp. 397-403, (June 1964).

8. Krasovskii, N. N., Stability of Motion, Stanford University Press, Stanford, California, (1963).

9. Lakshmikantham, V., "Liapunov Function and a Basic Inequality in Delay-Differential Equations," Archive for Rational Mechanics and Analysis, 10, pp. 305-310, (1962).

10. LaSalle, Joseph, and Lefschetz, Solomon, Stability by Liapunov's Direct Method With Applications, The RAND Corporation, Mathematics in Science and Engineering Series, 4, Academic Press, (1961).

11. Mohler, Ronald, R., "Optimal Control of Nuclear Reactor Processes," Doctoral Dissertation, University of Michigan, (To appear).

12. Monroe, Alfred J., Digital Processes for Sampled Data System, John Wiley and Sons, Inc., (1962).

13. Razumikhin, B. S., "The Application of Liapunov's Method to Problems in the Stability of Systems With Delay," Automation and Remote Control, 21, pp. 515-520, (1960).

14. _____, "Stability in First Approximation of Systems with Lag," Prikl. Mat. Meh., 22, pp. 155-166, English Translation, pp. 215-229, (1958).

15. Schultz, D. G., "The Variable Gradient Method of Generating Liapunov Functions With Applications to Automatic Control Systems," Doctoral Dissertation, School of Electrical Engineering, Purdue University, Lafayette, Indiana, (1962).

16. Seale, R. L., "Coupled Core Reactors," Los Alamos Scientific Laboratory Report No. LAMS-2967, (1964).

17. _____, "Investigation of Coupled ROVER Cores," Los Alamos Scientific Laboratory Report No. LAMS-2968, (1964).

18. Weaver, Lynn E., System Analysis of Nuclear Reactor Dynamics, American Nuclear Society and U. S. Atomic Energy Commission Monograph Series on Nuclear Science and Technology, Rowman and Littlefield, Inc., New York, (1963).

SECTION III

SYNTHESIS OF OPTIMAL CLOSED-LOOP CONTROL FOR NUCLEAR ROCKET SYSTEMS

Chapter 1

SYNTHESIS OF OPTIMAL CLOSED LOOP CONTROL FOR NUCLEAR REACTOR SYSTEMS

Theory

A number of papers in recent years have discussed the determination of an optimal control for nuclear reactor systems[1,2,3,4]. The optimal input control vector and the optimal output state vector are computed using Pontryagins's Maximum Principal[5] for the given performance criteria. The resulting control system typically operates open loop and thus will be quite susceptible to any internal noise or errors due to the non-exact mathematical description of the system. In order to reduce the effect of these disturbances on the optimum system state-transition trajectories, it is desirable to determine a closed loop controller.

The familiar neutron kinetics equations of a reactor system, which in general are nonlinear, are described in state variable notation by first order differential equations of the form

$$\dot{\bar{x}} = \bar{f}(\bar{x}, \bar{u}) \tag{1.1}$$

where $\bar{x}$ is the state vector and $\bar{u}$ the control vector. When random disturbances are added to the system the problem of determining the optimum control then becomes one of a statistical nature.

For linear dynamic systems, the well known Wiener filter was developed by Wiener to handle such a statistical problem. Because the

Wiener problem is solved in the frequency domain, a number of limitations curtail its usefulness.

(1) The optimal filter is specified by its impulse response, and the task of synthesizing the filter from such data is not easy.

(2) Numerical determination of the optimal impulse response is often quite involved making machine computation difficult.

(3) Important generalizations require new derivations which may be difficult to the nonspecialist.

R. E. Kalman and R. S. Bucy[7,8] in recent years have taken the Wiener filter problem in its entirety out of the frequency domain, restating it in the time domain. This new approach to linear filtering has virtually eliminated the major limitations associated with the Wiener approach, making synthesis by machine computation both feasible and relatively simple. This fact is demonstrated by a number of problems to which this approach has been applied since the appearance of references (7) and (8)[9,10,11,12]. The theory has been labled the "Linear Optimal Stochastic Control Theory."

As stated previously, the nuclear reactor systems considered in this study are described by sets of nonlinear first order differential equations. In order to apply linear optimal stochastic control theory, the equations must be made compatible with the theory. This is accomplished by making piecewise linear approximations about the predetermined optimal trajectories. The resulting piecewise linear differential equations describing perturbations about the optimum path are time varying and take the form

$$\frac{\bullet}{\delta \bar{x}} = \left[\frac{\partial \bar{f}}{\partial \bar{x}}\right]_{\bar{x}_o} \bar{\delta x} + \left[\frac{\partial \bar{f}}{\partial \bar{u}}\right]_{\bar{u}_o} \bar{\delta u} \qquad (1.2)$$

where $\overline{\delta x}$ is a small perturbation about the optimal output vector $\bar{x}_o$

$$\text{ie} \qquad \overline{\delta x} = \bar{x} - \bar{x}_o$$

and $\overline{\delta u}$ is a small perturbation about the optimal input control vector $\bar{u}_o$

$$\text{ie} \qquad \overline{\delta u} = \bar{u} - \bar{u}_o$$

For convenience Eq. (1.2) is redefined as

$$\dot{\overline{\delta x}} = F(t)\,\overline{\delta x} + G(t)\overline{\delta u} \tag{1.3}$$

where F(t) is the system matrix and G(t) is the input matrix.

In order to determine the optimal feedback control, one must first choose some index of performance to extremize. For this work quadratic indices of performances of the type

$$\delta J = \frac{1}{2}\,\overline{\delta x}^T(t_f)S(t_f)\overline{\delta x}(t_f) + \frac{1}{2}\int_o^{t_f} (\overline{\delta x}^T Q_1(t)\overline{\delta x} + \overline{\delta u}^T Q_2(t)\overline{\delta u})\,dt \tag{1.4}$$

were assumed where $S(t_f)$ is the terminal condition matrix whose elements are chosen to obtain the desired terminal accuracy; $Q_1(t)$ is the state variable error weighting matrix; $Q_2(t)$ is the control weighting matrix, and t_f is the terminal time.

Pontryagin's Maximum Principle is now applied to determine the optimal control feedback. The pre-Hamiltonian for the system becomes

$$R = \frac{1}{2}\,\overline{\delta x}^T Q_1(t)\overline{\delta x} + \frac{1}{2}\,\overline{\delta u}^T Q_2(t)\overline{\delta u} + \bar{\psi}^T(\dot{\bar{x}} - \bar{f}(\bar{x},\bar{u})_o)$$

$$= \frac{1}{2}\,\overline{\delta x}^T Q_1(t)\overline{\delta x} + \frac{1}{2}\,\overline{\delta u}^T Q_2(t)\overline{\delta u} + \bar{\psi}^T\left[F(t)\overline{\delta x} + G(t)\overline{\delta u}\right]$$

where $\psi(t)$ is the costate vector which is adjoint to the state variable

vector. Hamilton's equations for the system are

$$\dot{\overline{\delta x}} = \frac{\partial R}{\partial \overline{\psi}} = F(t)\overline{\delta x} + G(t)\overline{\delta u} \tag{1.6}$$

$$\dot{\overline{\psi}} = \frac{-\partial R}{\partial \overline{\delta x}} = -Q_1(t)\overline{\delta x} - F(t)^T\overline{\psi} \tag{1.7}$$

The optimal feedback control $\overline{\delta u}_o$ is that control which minimizes R (this is the system Hamiltonian)

$$H = 0 = \frac{\partial R}{\partial \overline{\delta u}} = Q_2(t)\ \overline{\delta u}_o + G^T(t)\ \overline{\psi} \tag{1.8}$$

solving for $\overline{\delta u}_o$

$$\overline{\delta u}_o = -Q_2^{-1}(t)G^T(t)\overline{\psi} \tag{1.9}$$

It is now assumed that the costate vector $\overline{\psi}$ is of the form

$$\overline{\psi} = P(t)\overline{\delta x} \tag{1.10}$$

Substituting Eq. (1.10) into Eq. (1.9) it is clear that the optimum control is

$$\overline{\delta u}_o = -Q_2^{-1}(t)G^T(t)P(t)\overline{\delta x} \tag{1.11}$$

where $-Q_2^{-1}G^TP$ is the feedback gain. Differentiating Eq. (1.10) with respect to time and employing Eqs. (1.6), (1.7), and (1.11), one can determine the form of P(t).

$$\dot{\overline{\psi}} = \dot{P}(t)\overline{\delta x} + P(t)\dot{\overline{\delta x}} = -Q_1(t)\overline{\delta x} - F^T(t)\overline{\psi}$$

$$\dot{P}(t)\overline{\delta x} + P(t)\left[F(t)\overline{\delta x} + G(t)\overline{\delta u}\right] = -Q_1(t)\overline{\delta x} - F^T(t)P(t)\overline{\delta x}$$

$$\dot{P}(t)\overline{\delta x} + P(t)F(t)\overline{\delta x} + P(t)G(t)\left[-Q_2^{-1}(t)G^T(t)P(t)\right]\overline{\delta x} = -Q_1(t)\overline{\delta x} - F^T(t)P(t)\overline{\delta x} \tag{1.12}$$

Since $\overline{\delta x}$ is common to all terms and is not equal to zero

$$\dot{P}(t) = -P(t)F(t) - F^{T}(t)P(t) - Q_1(t) + P(t)G(t)Q_2^{-1}(t)G(t)P(t) \tag{1.13}$$

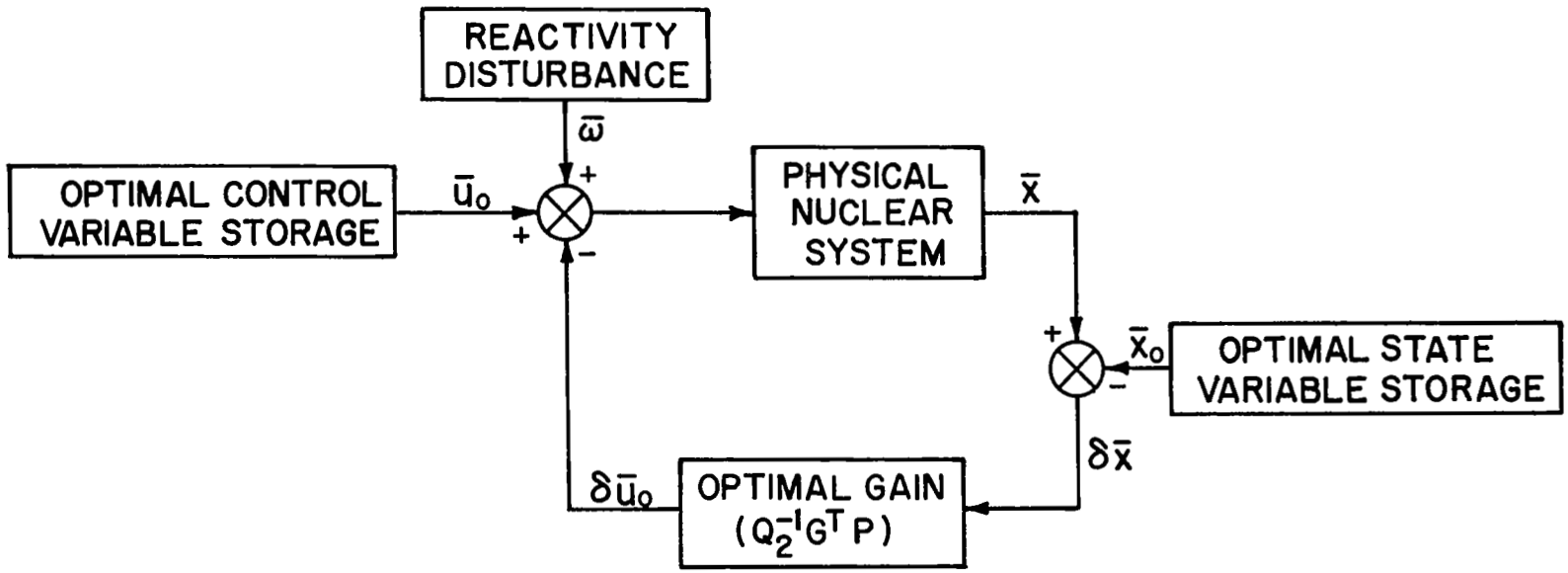
Here P(t) is a symmetric, nonlinear, generally time-varying matrix satisfying a differential equation of the matrix Ricatti type. If the terminal time is not permitted to vary, ie $\delta t_f = 0$, then the boundary condition on P(t) is found to be $P(t_f) = S(t_f)$.

Thus in order to solve for P(t) it is necessary to integrate Eq. (1.13) backwards in time.

Figure 1 is a block diagram of the solution of the optimal control process.

If, in addition to system internal noise, there is also noise in the measurement of the system state variables, which is generally the case, statistical estimates of these state variables must be made to predict the optimal control. The optimal closed loop control is then defined as that control which minimizes both internal and measurement noise. Since the optimal closed loop control, as shown, is based on known values of the system's state variables, it is important that the best estimate possible be determined.

It is known that for linear systems with quadratic performance criteria, it is possible to solve the estimation problem and the previously developed optimization problem separately and still obtain the overall optimum system[13]. When estimates of the state variables are used in nonlinear systems, however, it cannot be assumed that the overall system will still be optimal. Clearly, if estimates of the state variables were



SOLUTION OF THE OPTIMAL CONTROL PROCESS IN THE PRESENCE OF RANDOM REACTIVITY DISTURBANCES.

FIG. 1

available the only recourse would be to use them. The problem of joint estimation and optimization for a nonlinear system is an extremely difficult unsolved class of statistical optimization problems.

The differential equations governing the dynamics of the nuclear reactor system now become

$$\begin{aligned} \dot{\bar{x}} &= \bar{f}(\bar{x}, \bar{u}) + \bar{w}(t) \\ \bar{y} &= M(t)\bar{x} + \bar{v}(t) \end{aligned} \tag{1.14}$$

where

$\bar{x}(t)$ = state vector

$\bar{u}(t)$ = control vector

$\bar{y}(t)$ = measurable vector

$\bar{w}(t)$ = Markov-Gauss random vector

$\bar{v}(t)$ = Markov-Gauss random vector

The vectors $\bar{w}(t)$ and $\bar{v}(t)$ in Eq. (1.14) are independent random processes (white noise) with identically zero means and covariance matrices

$$\begin{aligned} \text{cov}\,[\bar{w}(t), \bar{w}(\tau)] &= A(t) \cdot \delta(t - \tau) \\ \text{cov}\,[\bar{v}(t), \bar{v}(\tau)] &= R(t) \cdot \delta(t - \tau) \quad \text{for all } t, \tau \\ \text{cov}\,[\bar{w}(t), \bar{v}(\tau)] &= 0 \end{aligned} \tag{1.15}$$

where $\delta(\cdot)$ is the Dirac delta function; $A(t)$ is a positive semidefinite symmetric matrix; $R(t)$ is a positive definite symmetric matrix.

The covariance matrix of two vector valued random variables $\bar{a}(t)$, $\bar{b}(\tau)$ is denoted by

$$\text{cov}\,[\bar{a}(t), \bar{b}(\tau)] = \mathcal{E}\,\bar{a}(t)\bar{b}^{-T}(\tau) - \mathcal{E}a(t)\,\mathcal{E}b^{T}(\tau) \tag{1.16}$$

where $\mathcal{E}(\cdot)$ denotes expected value.

In order to use linear optimal stochastic control theory, it is necessary to expand Eq. (1.14) about the optimal trajectory

$$\dot{\overline{\delta x}} = \left[\frac{\partial \bar{f}}{\partial \bar{x}}\right]_{|\bar{x}_o} \overline{\delta x} + \left[\frac{\partial \bar{f}}{\partial \bar{u}}\right]_{|\bar{u}_o} \overline{\delta u} + \bar{w} \tag{1.17}$$

or

$$\dot{\overline{\delta x}} = F(t)\overline{\delta x} + G(t)\overline{\delta u} + \bar{w}$$

$$\overline{\delta y} = M(t)\overline{\delta x} + \bar{v}$$

Here it is assumed that perturbations of the state variables about their optimal paths are due entirely to noise in the control vector. In a nuclear reactor system this is reasonable because all noise will show up as reactivity perturbation which in most cases is the control variable.

The optimal estimate of the state vector at time t based on known information for time prior to t is $\hat{\overline{\delta x}}(t|t)$ and is generated by a linear dynamical system of the form

$$\dot{\hat{\overline{\delta x}}}(t/t) = F(t)\hat{\overline{\delta x}}(t/t) + K(t)\hat{\overline{\delta y}}(t/t)$$

$$\hat{\overline{\delta y}}(t/t) = \hat{\overline{\delta y}} - M(t)\hat{\delta x}(t/t) \tag{1.18}$$

The initial state $\overline{\delta x}(t_o|t_o)$ is zero.

The optimal filter required for state variable estimation is a feedback system. It is obtained by taking a linear model of the plant dynamics, omitting the input control, forming the error signal $\hat{\overline{\delta y}}(t|t)$ and feeding the error forward with a gain K(t). Thus the specification of the optimal estimation filter is the computation of the optimal time varying gain K(t).

The optimal error is denoted

$$\underset{\sim}{\hat{\delta x}}(t|t) = \underset{\sim}{\delta x} - \underset{\sim}{\hat{\delta x}}(t|t) \tag{1.19}$$

This optimal error $\underset{\sim}{\hat{\delta x}}(t|t)$ is governed by the same dynamics as $\underset{\sim}{\hat{\delta x}}(t|t)$

$$\dot{\underset{\sim}{\hat{\delta x}}}(t|t) = F(t)\underset{\sim}{\hat{\delta x}}(t|t) + G(t)\bar{\delta u} - K(t)\left[\bar{v} + M(t)\underset{\sim}{\hat{\delta x}}(t|t)\right] \tag{1.20}$$

where $\underset{\sim}{\hat{\delta x}}(t_o|t_o)$ is zero.

Kalman has derived the form of the optimal gain. This derivation is abstract and lengthy, and the reader is referred to sections 8,9, and 10 of reference (8) for details. The optimal gain is

$$K(t) = V(t)M^T(t)R^{-1}(t) \tag{1.21}$$

The matrix V(t) is the covariance matrix of $\underset{\sim}{\hat{\delta x}}(t|t)$

$$V(t) = \text{cov}(\underset{\sim}{\hat{\delta x}}(t|t), \underset{\sim}{\hat{\delta x}}(t|t)) \tag{1.22}$$

Kalman has also shown that V(t) must be the solution of the Ricatti-type matrix differential equation

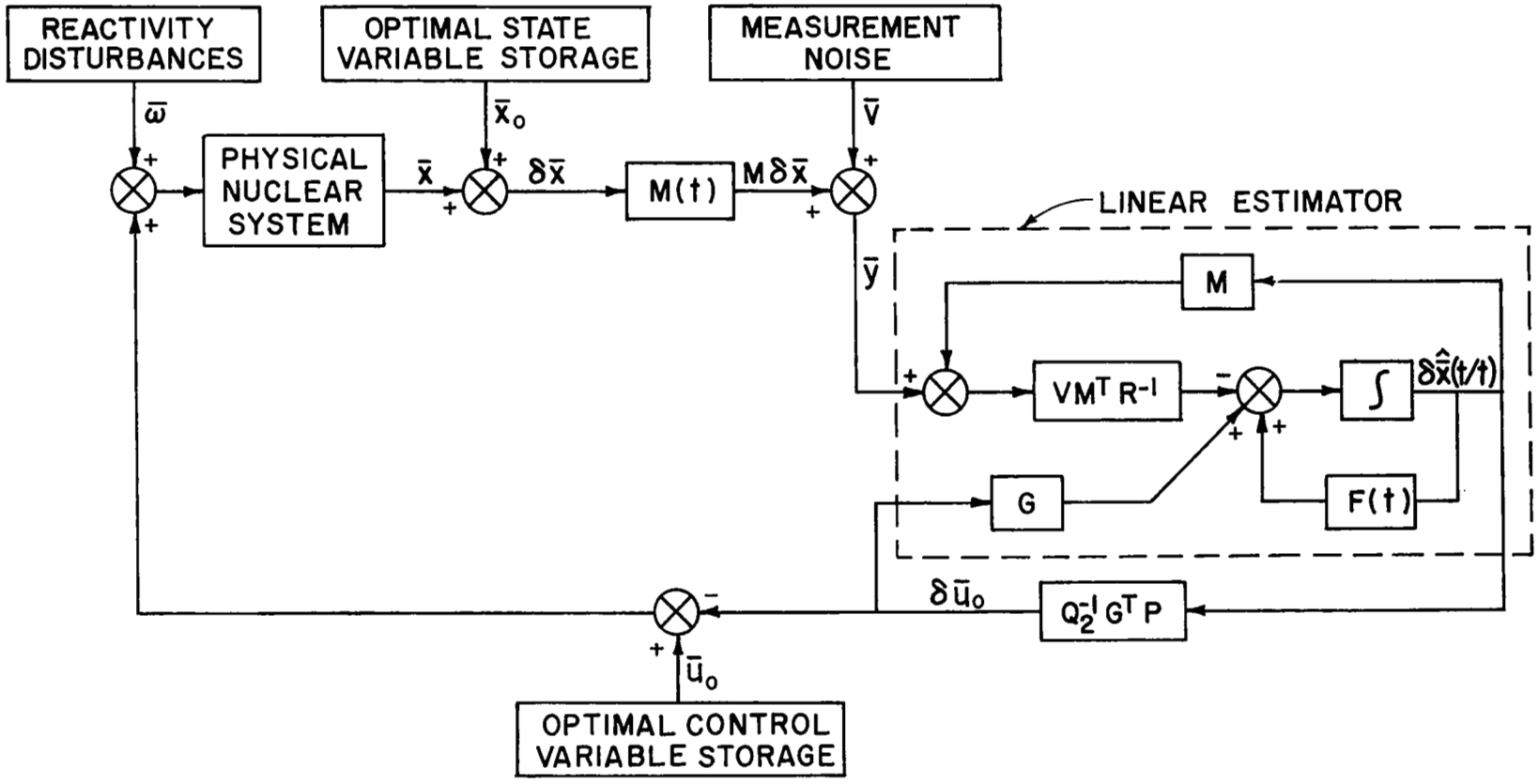
$$\dot{V}(t) = F(t)V(t) + V(t)F^T(t) - V(t)M(t)R^{-1}(t)M(t)V(t) + G(t)A(t)G^T(t) \tag{1.23}$$

The initial condition that must be satisfied for Eq. (1.22) is

$$V(t_o) = \text{cov}(\bar{\delta x}(t_o), \bar{\delta x}(t_o)) \tag{1.24}$$

Figure 2 is a general block diagram of the solution of the optimal estimation and control problem.

Since Kalman's formulation and solution of this problem of estimating the state variables, a great deal of interest has been shown by a number of other researchers.[14,15,16,17] Their work basically present



SOLUTION OF THE OPTIMAL CONTROL BY LINEAR ESTIMATION OF THE STATE VARIABLE RANDOM DISTURBANCES.

FIG. 2

somewhat simpler techiques for obtaining optimum estimates of the state variables for discrete time systems. Both the linear and the nonlinear estimation problems have been investigated with the unanimous result that, given all values of the measurable vector $\overline{\delta y}$ up to $\overline{\delta y}(t)$, one can statistically determine the optimal estimate $\hat{\overline{\delta x}}(t|t)$.

For the nonlinear problems undertaken in this study a modification of Kalman's linear approach was formulated for the optimal estimate. This modification partially relaxes the requirement that the optimal estimation of the system state variables be determined solely from a linearized model of system dynamics about the optimal trajectories. Upon examination of Eq. (1.17) it is seen that the dynamics of the optimal error is a linear function, of both the optimal error $\hat{\overline{\delta x}}(t|t)$ and the optimal feedback control $\overline{\delta u}(t)$ whereby it is formulated that the optimal estimate can be given by the equation

$$\begin{aligned} \dot{\overline{x}}(t|t) &= \overline{f}(\hat{\overline{x}}(t|t), \overline{u}(t)) + K(t)\hat{\overline{y}}(t|t) \\ \hat{\overline{y}}(t|t) &= \overline{y}(t) - M(t)\hat{\overline{x}}(t|t) \end{aligned} \tag{1.25}$$

where it is assumed that

$$\hat{\overline{x}}(t_o|t_o) = \overline{x}(t_o) = \overline{x}_o(t_o) \tag{1.26}$$

It is observed, however, that there is a descrepancy between Kalman's optimal estimate given by Eq. (1.18) and that formulated in Eq. (1.25). Eq. (1.25) is an explicit function of the control vector whereas Eq. (1.18) is independent of the control. The justification for using Eq. (1.25) rather than Eq. (1.18) can be shown by a simple illustration.

If it is assumed that there is no noise, either internally or in the measurement, it is obvious that the optimal estimate will be identical with the physical state defined by Eq. (1.1) in the nonlinear case and by Eq. (1.3) in the linear case.

nonlinear, $\hat{\bar{x}}(t|t) = \bar{x}(t)$ for $\bar{w} = 0$ and $\bar{v} = 0$

ie

or linear $\hat{\underset{\sim}{\delta\bar{x}}}(t|t) = \underset{\sim}{\delta\bar{x}}(t)$

Clearly for this condition

$$\text{a) } \hat{\delta\bar{y}}(t|t) = 0 \qquad \text{and} \qquad \text{b) } \hat{\bar{y}}(t|t) = 0 \tag{1.27}$$

Substituting conditions (1.27a) and (1.27b) into Eq. (1.19) and Eq. (1.25) respectively

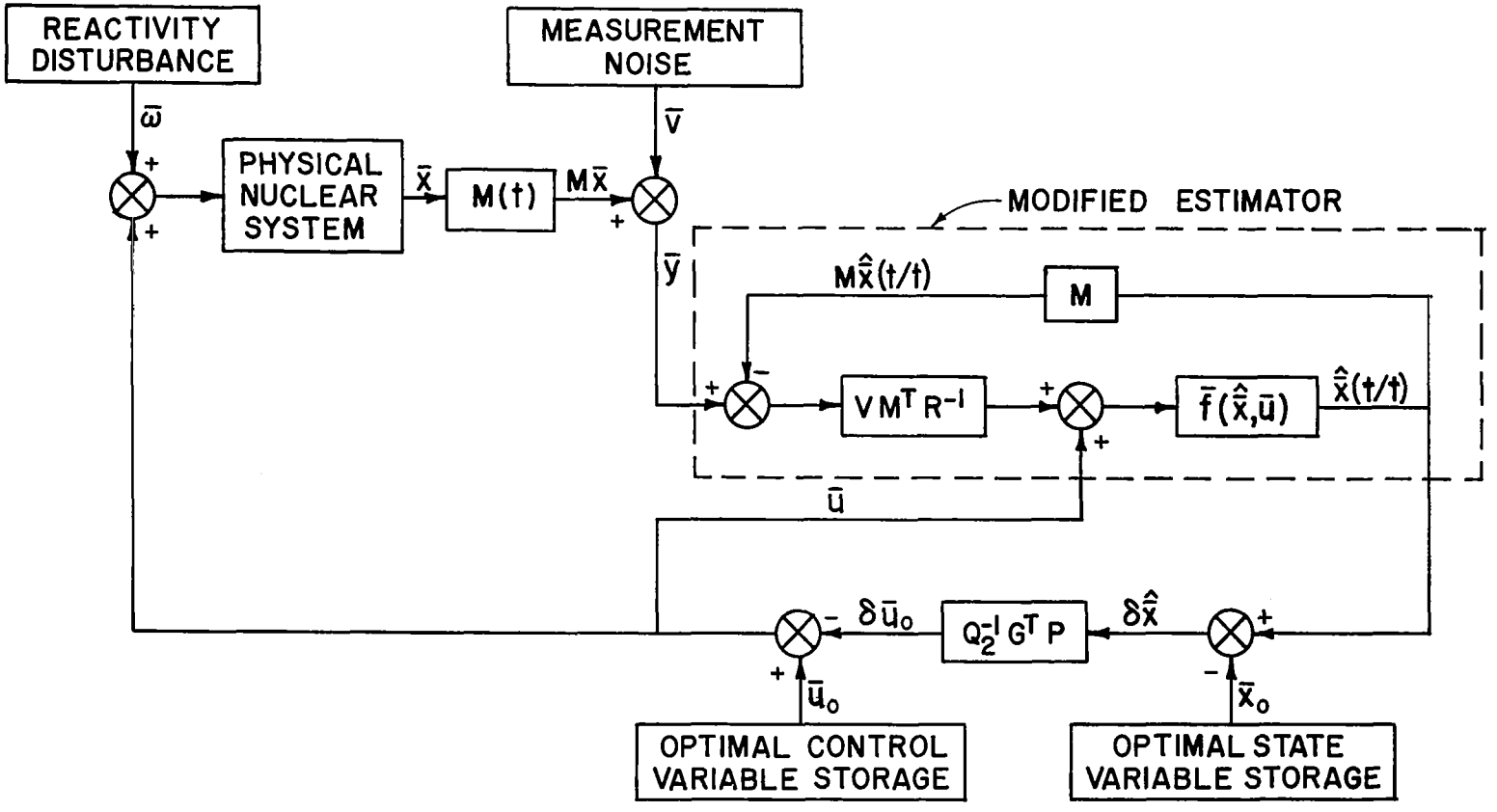
$$\dot{\delta\bar{x}}(t) = F(t)\delta\bar{x}(t) = 0 \tag{1.3a}$$

$$\dot{\bar{x}}(t) = \bar{f}(\bar{x}(t), \bar{u}(t)) \neq 0 \tag{1.1a}$$

Eq. (1.3a) is no longer a linear function of the control $\delta\bar{u}$ and thus $\delta\bar{x}$ can only be zero for this case. This is an obvious result since no perturbations exist. Eq. (1.1a), on the other hand, not only is a function of the control, but it is identical with Eq. (1.1).

This formulation lacks rigorous proof. It is used in this study solely on the basis of the above illustration. Figure 3 is a general block diagram of the solution of the optimal control and estimation problem just discussed.

The optimal feedforward gain K(t) is still determined by Eqs. (1.21), (1.22), and (1.24). Therefore, the dynamics are linearized only for the purpose of solving the matrix - Ricatti-type differential Eq. (1.23). This scheme for determining the optimal estimate works nicely for the class of nonlinear problems used to describe nuclear reactors.



SOLUTION OF THE OPTIMAL CONTROL BY NON-LINEAR ESTIMATION OF THE STATE VARIABLES WITH RANDOM DISTURBANCES.

FIG. 3

In general the nonlinearities are never more than the product of two system variables. There is no guarantee, however, that estimates of the state variables for dynamic systems with other types of nonlinearities can be successfully made using this technique. Higher order errors, products of erros, and/or division by errors may tend to obscure the estimates based on Eq. (1.25).

A result of the linear estimation which makes it an optimal approach, is that as time increases the statistics necessary to determine K(t) get progressively better. Since Eq. (1.18) is a superposition of only linear terms, the optimal estimate theoretically is identical with the actual physical state at infinite time. The same cannot be said about the superposition of a linear component with nonlinear components as in Eq. (1.25). In fact, in some other classes of nonlinear problems it may be that the statistics become worse. This, of course, would not only make successive estimates worse, but would increase the perturbations about the optimal state.

Chapter 2

OPTIMUM SYNTHESIS APPLIED TO NUCLEAR REACTOR POWER TRANSFER

This particular synthesis technique was applied to a problem of power state change in a TRIGA type nuclear reactor system. The optimum state transition trajectories were determined in reference 2.

The problem of interest considers a bare thermal reactor with temperature feedback. For convenience only one group of delayed neutrons were used. It is also assumed that the core temperature is proportional to the power level. The total effective system reactivity is then the sum of the external control rod reactivity input ρ and the temperature feedback reactivity.

$$\rho_t = \rho - \alpha n \tag{2.1}$$

where α is the power (temperature) coefficient of reactivity and n is the reactor power level. The reactor kinetics are described by the following equations.

$$\dot{n} = (\rho t-\beta)n/\ell + \lambda c = (\rho-\alpha n-\beta)/\ell + \lambda c$$

$$\dot{c} = \beta n/\ell - \lambda c \tag{2.2}$$

where the neutron density (power level) n and the precursor concentration c are the state variables and the reactivity is the control variable. The system is assumed to be in the steady-state for time $t \leq 0$ and has the initial conditions.

$$n(0) = n_o \ , \ c(0) = c_o \tag{2.3}$$

The problem is to increase the power from the initial state n_o to a terminal state an_o, where a is some constant greater than 1.0, with minimum control energy. The performance index for the system is

$$J = \int_o^{t_f} \dot{\rho}^2 \, dt \tag{2.4}$$

It is assumed that the control system has inertia and cannot respond instantaneously. This is given by the constraint

$$|\dot{\rho}| \leq \dot{\rho}_{max} \tag{2.5}$$

The control variable thus becomes $\dot{\rho}$ in lieu of ρ and ρ now becomes a state variable. The kinetics equations become

$$\begin{aligned} \dot{n} &= (\rho - \alpha n - \beta)n/\ell + \lambda c \\ \dot{c} &= \beta n/\ell - \lambda c \\ \dot{\rho} &= u \end{aligned} \tag{2.6}$$

where u is the control variable. Figures 4 and 5 show the optimal reactivity and optimal power level trajectories, respectively, for increasing the reactor power from 10 kw to 50 kw with minimum energy. Figure 6 shows the control variable optimal trajectory.

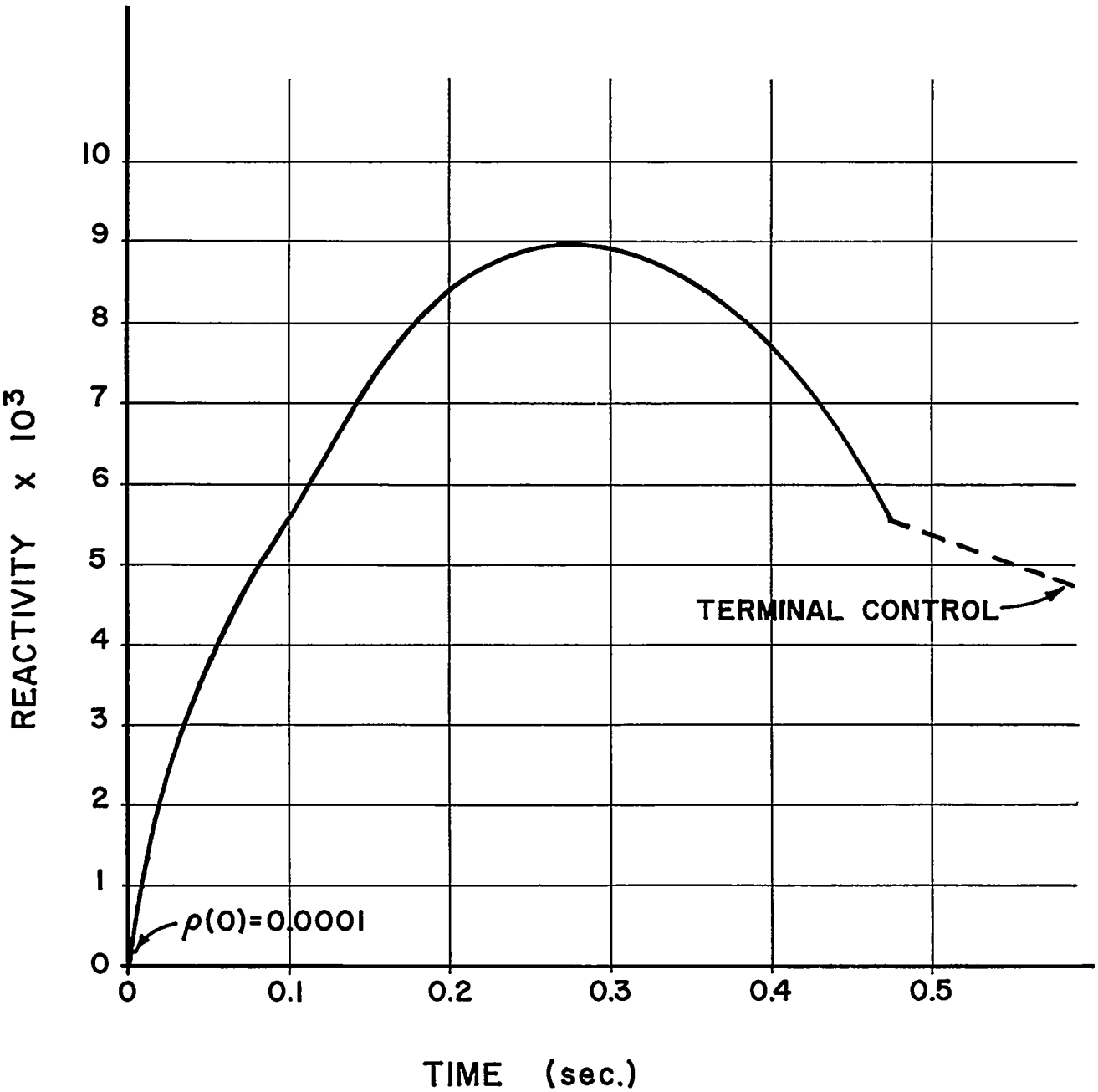
The problem stated above was reformulated for digital computer computation. The optimal power trajectory was approximated by an eighth degree polynomial with time as the variable. The following parameters were used for computation.

$\lambda = 0.1 \text{ sec}^{-1}$ $\qquad$ $n_o = 10$ kw

$\alpha = 10^{-5} \text{ kw}^{-1}$ $\qquad$ $a = 5$

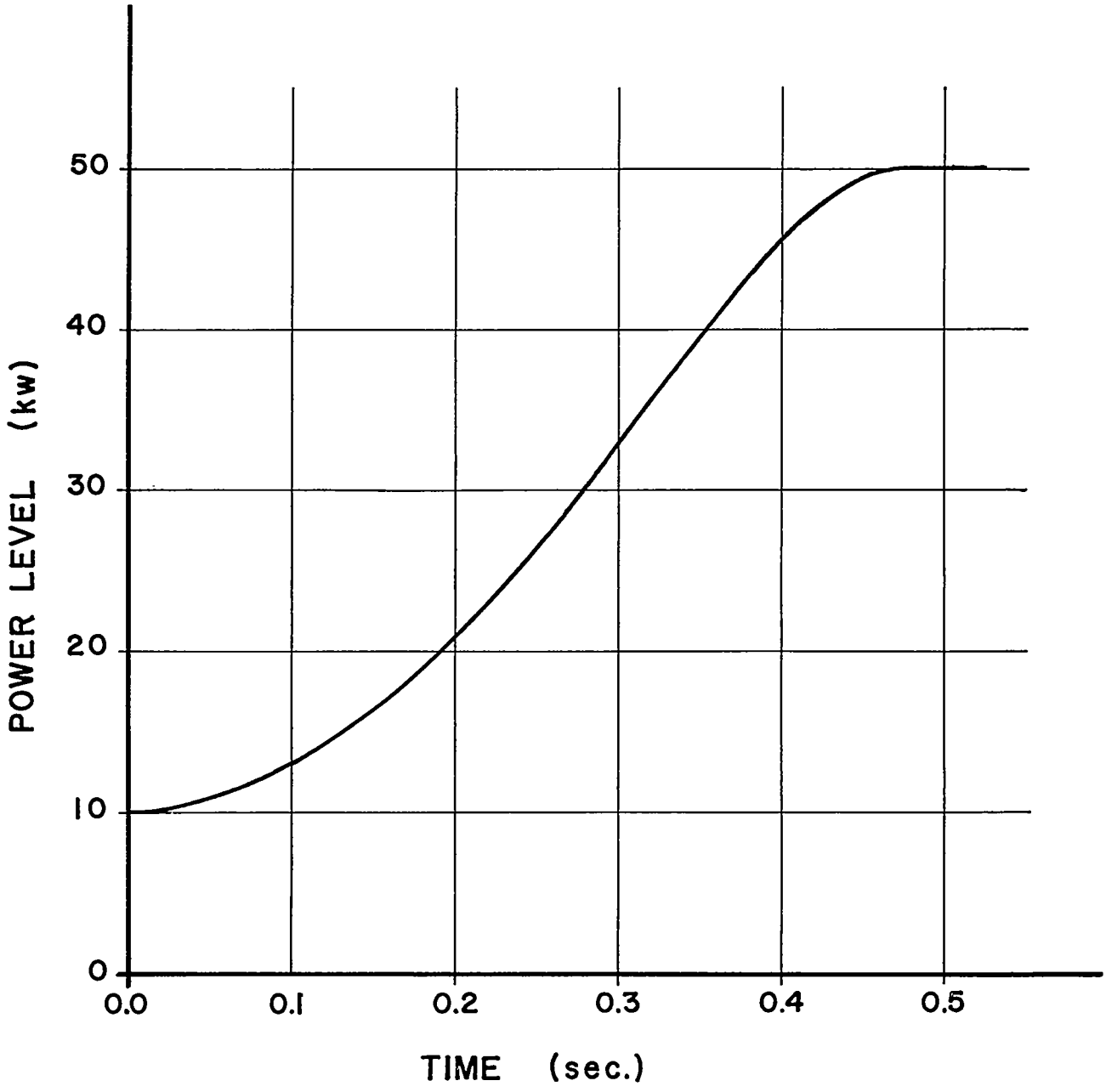
$\ell = 10^{-3}$ sec $\qquad$ $\beta = 0.0064$

$t_f = 0.47$ sec



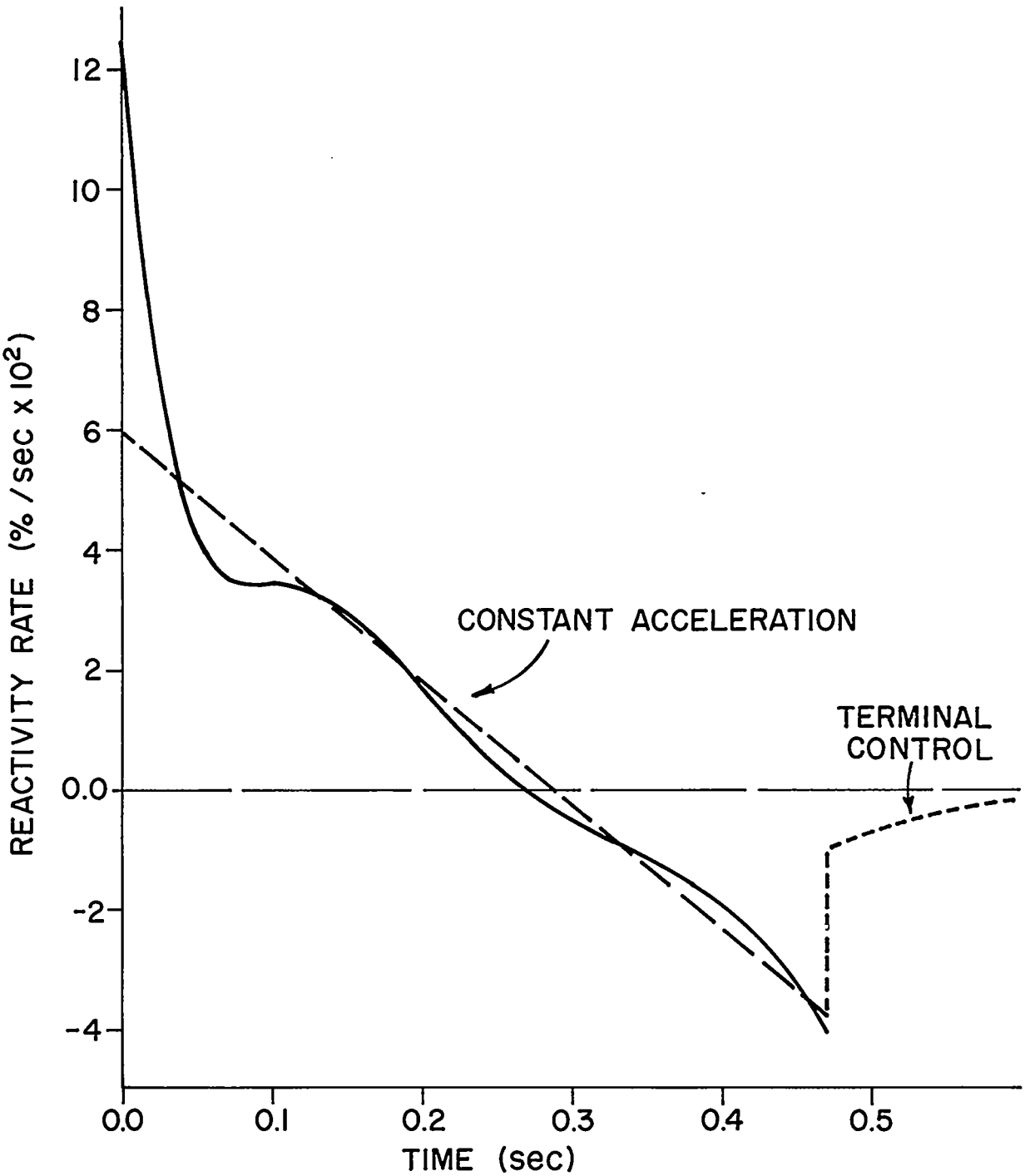
OPTIMAL RELATIVITY PROGRAM FOR CHANGE OF POWER LEVEL FROM 10 KW TO 50 KW ; MINIMUM CONTROL ENERGY.

FIG. 4



OPTIMAL POWER LEVEL PROGRAM FOR MINIMUM ENERGY CONTROL.

FIG. 5



OPTIMAL CONTROL PROGRAM FOR POWER LEVEL CHANGE
FROM 10kw TO 50kw
MINIMUM CONTROL ENERGY

FIGURE 6

The boundary condition at terminal time is that $n(t_f) = 0$. It is easily shown that the terminal reactivity required to maintain this condition is

$$\rho(t) = \ell\dot{c}(t)/n(t_f) + \alpha n(t_f), \quad t \geq t_f \tag{2.7}$$

For closed loop synthesis, this terminal reactivity can be maintained by a dither type control as suggested in reference 2 with appropriate reactivity constraints. This then requires that the optimal closed loop process only be used up to the terminal time t_f when $\dot{n}(t_f) = 0$ and $n(t_f) = an_o$.

Having formulated a model for the nuclear reactor system and determined the optimal trajectories it is necessary to determine the linearized system coefficient matrix about the optimal trajectories. The deviation of the state and control variables about the optimal trajectories are

$$\left.\begin{aligned} \delta n &= n_{ac} - n_{op} \\ \delta c &= c_{ac} - c_{op} \\ \delta\rho &= \rho_{ac} - \rho_{op} \end{aligned}\right\} \text{ State Variables} \tag{2.8}$$

and

$$\delta u = u_{ac} - u_{op} \tag{2.9}$$

The output state vector is

$$\overline{\delta x} = \begin{bmatrix} \delta n \\ \delta c \\ \delta\rho \end{bmatrix} \tag{2.10}$$

The linearized model becomes

$$\dot{\overline{\delta x}} = \begin{bmatrix} f_{11}(t) & \lambda & f_{13}(t) \\ \beta/\ell & -\lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \overline{\delta x} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \delta u \tag{2.11}$$

$$= F(t)\overline{\delta x} + G(t)\overline{\delta u} \tag{2.12}$$

where

$$f_{11}(t) = (\rho_{op} - \alpha n_{op} - \beta)/\ell \quad .$$

$$f_{13}(t) = n_{op}/\ell \qquad (2.13)$$

The steps involved in the realization of the control system are:

(a) Choice of the acceptable optimal trajectories.

(b) Evaluation of F(t) and G(t) along the optimal trajectories.

(c) Solution of P(t) and V(t) via the matrix Ricatti-type Eqs. (1.13 and (1.22).

(d) Storage of the optimal state variables, optimal control variables and the feedback gains $Q_2^{-1}G^TP$ and VM^TR^{-1}.

The selection of the error weighting and control weighting matricies, Q_1 and Q_2, is arbitrary. In practice the best selection is determined by varying the ratio of Q_1/Q_2. Theoretically the optimum feedback control system is arrived when this ratio becomes infinite. In actuality, however, this situation is physically unrealizable since the control system would become so "sluggish" that it could not possibly follow the system. If the ratio is too small the feedback control would not have much effect on the random process. The obvious choice, then is the smallest ratio which yields acceptable accuracy. The control weighting matrix Q_2 for this problem is a 1x1 matrix equal to unity. From Eq. (1.11) the optimal feedback control becomes

$$\delta u_o = -(P_{31}\widehat{\delta n} + P_{32}\widehat{\delta c} + P_{33}\widehat{\delta\rho}) \qquad (2.14)$$

where P_{ij} are typical elements of the P(t) matrix solution of Eq. (1.13). The error weighting matrix is a 3 x 3 diagonal matrix of the form

$$Q_1 = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix} \qquad (2.15)$$

where the elements A, B, and C are chosen to obtain the required accuracy. For convenience A, B, and C are constant and chosen to be integral powers of 10. The matrix can easily be represented in the following way,

$$Q_1 = Q_1(a, b, c) = \begin{bmatrix} 10^a & 0 & 0 \\ 0 & 10^b & 0 \\ 0 & 0 & 10^c \end{bmatrix} \tag{2.16}$$

where a, b, c are the exponents.

Figure 7 illustrates feedback gain programs of P_{31}, P_{32}, and P_{33} for continuous error detection in this example, for Q_1 $(0, -\infty, 4)$. This solution was obtained by integrating Eq. (1.13) backwards in time on The University of Arizona's IBM 7072 digital computer. The techniques required to obtain this solution are discussed fully in the section on numerical methods.

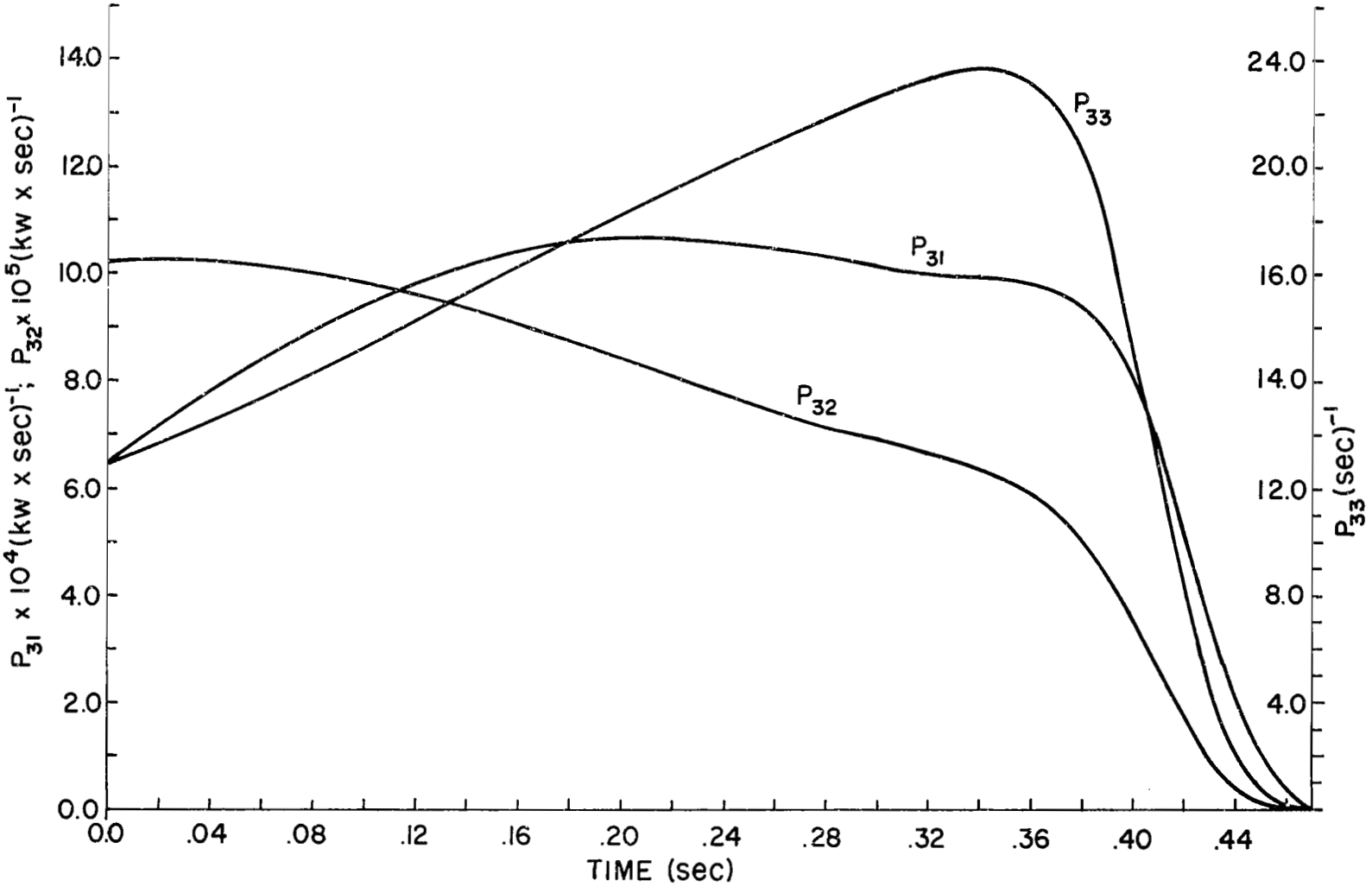
As mentioned, random disturbances are incorporated into the nuclear reactor systems as control perturbations. For digital computer synthesis this is very easy to accomplish. The actual control becomes

$$u_{ac} = u_{op} + w - \delta u_{op} \tag{2.17}$$

where w is the random disturbance. This disturbance is assumed to be Gaussian with zero-mean. On the digital computer a random number generator is used to generate w. A discussion of the computer scheme is given later.

The covariance matrix $A(t)$ can be determined from w. Here w is a 1 x 1 vector with element w_{11}. The covariance matrix is then simply

$$A(t) = a_{11} = \int_{-\infty}^{\infty} \frac{w_{11}^2 \, e^{-w_{11}^2/2\sigma w_{11}^2}}{\sqrt{2\pi}\ \sigma w_{11}} \, dw_{11} = \sigma w_{11}^2 \tag{2.18}$$



FEEDBACK GAIN PROGRAMS FOR CONTINUOUS TIME ERROR DETECTION OF SYSTEM STATE VARIABLES FOR ERROR WEIGHTING MATRIX $Q_2(0, -\infty, 4)$

FIGURE 7

where σw_{11}^2 is the variance associated with the Gaussianly-distributed disturbance w_{11}. Inasmuch as synthesis of this problem was performed on a digital computer the statistics of w_{11} and other random variables were predetermined. By digital computer synthesis it is possible to determine how small the perturbations must be in order that this linearized feedback formulation be valid, by varying the statistics of the random variables.

In a reactor system it is only possible to measure the state variables of power and reactivity. No measurement can be performed on the precursor concentration. The measurement matrix must then take the form

$$M(t) = \begin{bmatrix} m_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & m_{33} \end{bmatrix} \tag{2.19}$$

For this example it can be assumed that the measurement is a linear function of the observed variables, ie $m_{11} = m_{33} = 1$. Frequently the power level of a large reactor is measured logarithmically changing the form of m_{11}. Noise in the measurement is represented by the Markov-Gauss vector

$$\bar{V} = \begin{bmatrix} Vn \\ Vc \\ V\rho \end{bmatrix} \tag{2.20}$$

The covariance matrix R(t) is determined for the case where all Vi have zero mean with no cross correlation

$$R(t) = \begin{bmatrix} \sigma n^2 & 0 & 0 \\ 0 & \sigma c^2 & 0 \\ 0 & 0 & \sigma \rho^2 \end{bmatrix} \tag{2.21}$$

The following values have been determined for the variances of the random variables. Here it is assumed that the random variables be

constrained to five per cent of the maximum value of their respective state or control variable 99.9 per cent of the time.

$$\sigma w^2 = 0.0000152$$

$$\sigma n^2 = 0.583$$

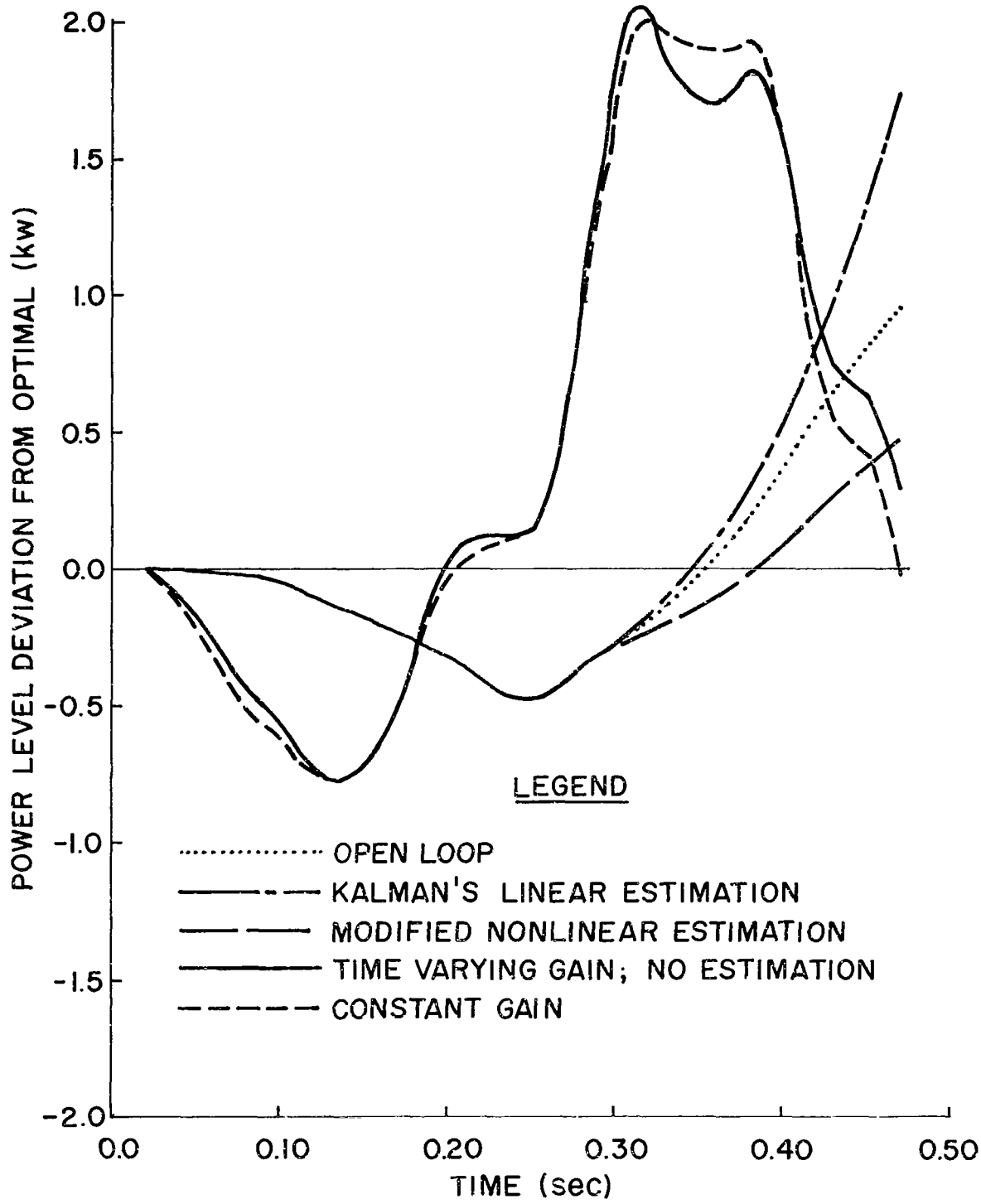
$$\sigma c^2 = 114.3$$

$$\sigma \rho^2 = 0.000168$$

Knowing the forms of M(t), A(t), and R(t) solution of Eq. (1.22) is possible. The elements of the covariance matrix V(t) are stored as the optimal filter gains. The same numerical techniques required to evaluate the P(t) matrix hold in this case with the exception that the V_i are integrated forward in time.

In practice time varying feedback gains are not desirable. It has been suggested that the matrix Ricatti differential equations be solved in the steady-state for linear systems to obtain constant gains. This technique, however, can not be applied to nonlinear systems due to the fact that both F(t) and G(t) are time varying functions of the optimal trajectories. A simple average of the time varying gains over the control period would accomplish the same purpose. However, such a control system would be sub-optimal since this would place too much control in the initial phase and too little control in the final phase.

Several control schemes were synthesized for this problem to determine the relative merits of each. Each scheme was carried out digitally using identical disturbances in both control and measurement. In all cases solution of the P(t) matrix was required. In addition the open loop solution was obtained for comparison purposes. Figure 8 illustrates the different power level trajectories for the four schemes



COMPARISON OF 4 OPTIMAL CLOSED LOOP CONTROL SCHEMES WITH OPEN LOOP AND OPTIMAL TRAJECTORIES.

FIGURE 8

with the inclusion of both the open loop and the optimal trajectories. These curves were obtained for an error weighting matrix of $Q_1(0, -\infty,4)$. For purposes of identification these schemes were labeled:

(1) Kalman linear estimator

(2) Nonlinear estimator

(3) Time varying gain with no estimator

(4) Constant gain with no estimator

The "Kalman linear estimator" is depicted in Figure 2. The solution of the "nonlinear estimator" is shown in Figure 3. These two systems require the solution of the matrix Ricatti-type Eq. (1.22). In practice this is done in real time during the control process using all known information. This, then, clearly requires in addition, not only stored programs of the optimal trajectories and gains, but a computing device to perform the integration. In most nuclear control applications such a system is not feasible. For controlled startup of a space nuclear reactor, however, where control times are small and optimization of some performance index, such a propellant consumption, is critical such control schemes appear desirable. Solution of the P(t) matrix must be precalculated as mentioned.

The "time varying gain with no estimator" scheme eliminates the requirement that a computing device be part of the control system. The solution of this scheme is shown in Figure 1. No attempt has been made to get statistically better estimates of the measured state variables. Since it is impossible to measure precursor concentration, and no attempt is made to estimate this variable, the form of the optimal control

changes from that given in Eq. (2.14) to

$$\delta u_{op} = -(P_{31}\widehat{\delta n} + P_{33}\widehat{\delta \rho}) \tag{2.22}$$

A sufficient error weighting matrix for this control scheme would be

$$Q_1 = Q_1(a, -\infty, b) \tag{2.23}$$

neglecting the weight on the unobservable state variable, precursor concentration.

The "constant gain with no estimator" is identical with the previous scheme except that the gain elements are averaged over the control period. This is the simplest control that still contains an element of the optimization technique. For most practical applications this would appear to be the most desirable. Solution of the matrix Ricatti-type Eq. (1.13) can easily be obtained for any desired trajectory. Here it should be pointed out that optimal closed loop control theory can be applied to any known nominal trajectory and not necessarily an optimal one, since only the error of the control and state variables are optimized.

Chapter 3

FORMULATION OF THE OPTIMAL CLOSED LOOP CONTROL PROBLEM FOR START UP OF A NUCLEAR ROCKET ENGINE

The dynamics of a nuclear rocket engine have been formulated in several references[18,19,20]. Two, somewhat conflicting, sets of nonlinear differential equations have been formulated by Smith and Stenning[18], and Mohler and Perry[19]. Both formulations consider a nuclear rocket engine with bleed turbo-pump or topping turbo-pump drive. The basic difference in these two concepts arises in the form of the temperature reactivity, δK_T. Smith and Stenning contend that this reactivity is directly proportional to the square root of the core exit stagnation temperature, ie $\delta K_T = \alpha_T \sqrt{T}$. Mohler and Perry contend that this reactivity is directly proportional to this temperature, ie $\delta K_T = \alpha'_T T$. Since there are other sources of reactivity in such a system both contentions could give fairly accurate results simply by choosing appropriate reactivity coefficients.

The model considered for this problem is that put forth by Mohler and Perry[19]. It consists of the basic neutron kinetics equations, coupled with a heat exchange equation via core temperature and propellant flow rate in the form of reactivity. The following nonlinear differential equations describe the system of interest.

Neutronics

$$\dot{Q} = \frac{\varrho_T - \beta}{\ell} Q + \sum_{i=1}^{6} \lambda_i C_i$$

$$\dot{C}_i = \beta_i Q/\ell - \lambda_i C_i \qquad (3.1)$$

Heat Exchanger $$\dot{T} = \frac{Q}{M} - \frac{T}{\tau_h} \tag{3.2}$$

where the system state variables are Q, C_i, and T.

Q = power level

C_i = precursor density for ith delayed neutron group

T = core exit stagnation temperature

ρ_T = total reactivity

M_c = mean effective heat capacity of reactor core

τ_h = heat exchanger time constant.

The total reactivity is comprised of control rod reactivity u_1, propellant density reactivity δK_p and temperature reactivity δK_T.

$$\rho_T = u_1 + \delta K_T + \delta K_p \tag{3.3}$$

where

$$\delta K_T = C_T T \tag{3.4}$$

and

$$\delta K_p = C_p u_2 / \sqrt{T} \tag{3.5}$$

Here u_1, is the control rod reactivity, a control variable and u_2 is the coolant mass flow rate, also a control variable. Usually $C_T \leq 0$ and $C_p \geq 0$. The heat exchanger thermal time constant is

$$\tau_h = (au_2)^1 \text{ sec} \tag{3.6}$$

where a is a constant of proportionality at rated design flow rate.

The following hypothetical nuclear rocket rated design conditions were used as system parameters.

Maximum Reactor Power, Q_{max}	2260 megawatts
Design Propellant Flow Rate	130 lb/sec

Maximum core exit stagnation temperature, T_{max}	4500 ^{o}R
Heat exchanger thermal time constant, τ_h	1.5 sec
Propellant inlet temperature, T_{min}	120 ^{o}R
Mean effective neutron lifetime, ℓ	3×10^{-5} sec
Propellant reactivity, δK_{P_o}	0.0065
Temperature reactivity, δK_{T_o}	-0.0065
Effective delayed neutron fraction, β	0.0065
Effective one group decay constant, λ	0.1 sec^{-1}
Effective core mass heat capacity, M_c	1140 Btu/^{o}R

The optimal control problem was stated by Mohler as follows:[21]

> "Given an initial reactor steady-state, bring the system to the desired terminal steady-state so as to minimize the consumption of propellant...."

The index of performance thus becomes the minimization of the control variable u_2. Due to turbo-pump design constraints, such as stalling and pump cavitation, u_2 is constrained to both an upper and lower limit.

$$u_a \leq u_2 \leq u_b \tag{3.7}$$

In addition to flow rate constraints there are constraints on core maximum power, maximum temperature, maximum and minimum control rod reactivity insertion, and rate of core temperature rise.

$$Q \leq Q_{max}$$

$$T \leq T_{max}$$

$$-\gamma\beta \leq u_1 \leq \gamma\beta$$

$$\dot{T}_{max} = \alpha$$

where γ is a positive number greater than 1.

For convenience, one group of delayed neutrons were used in this formulation. The pre-Hamiltonian for the system becomes:

$$\begin{aligned} R &= Q\psi_1 + C\psi_2 + T\psi_3 + u_2\psi_4 \\ &= I(u_1,Q,T,C) + \frac{Q\rho_T}{\ell}\psi_1 + u_2(\psi_4 - T_a\psi_3) \end{aligned} \tag{3.9}$$

The Hamiltonian is that function which minimizes R with respect to the control u_2

$$H = \frac{\partial R}{\partial u_2} = 0 = \frac{Q\rho T}{\ell}\psi_1 - Ta\psi_3 + \psi_4 \tag{3.10}$$

The optimal control scheme which satisfies all of the constraints and the index of performance has been determined by Mohler and is that used for this formulation. The control law was formulated digitally using Eqs. (3.1) and (3.2) and constraints (3.8). However, a fundamental discrepancy was encountered. When maximum power and maximum rate of temperature rise were achieved, both control variables u_1, and u_2 behaved opposite to that expected. The rod reactivity u_1, which should decrease at maximum power, increased monotonically. The propellant flow rate u_2, constrained to be minimum should have increased at the time $Q = Q_{max}$ and $\dot{T} = \alpha$, but decreased below the minimum. At present, the ambiguity has not been explained, The formulation of the closed loop control, however, still holds.

The deviation of the state and control variables about the optimal trajectories are:

$$\begin{aligned} \delta Q &= Q_{ac} - Q_o \\ \delta C &= C_{ac} - C_o \\ \delta T &= T_{ac} - T_o \end{aligned} \tag{3.11}$$

$$\begin{aligned} \delta u_1 &= u_{1ac} - u_{1o} \\ \delta u_2 &= u_{2ac} - u_{2o} \end{aligned} \tag{3.12}$$

The linearized output state vector is

$$\overline{\delta x} = \begin{bmatrix} \delta Q \\ \delta C \\ \delta T \end{bmatrix} \tag{3.13}$$

and the linearized control vector is

$$\overline{\delta u} = \begin{bmatrix} \delta u_1 \\ \delta u_2 \end{bmatrix} \tag{3.14}$$

The model describing perturbations about the optimal trajectories is:

$$\dot{\overline{\delta x}} = F(t)\overline{\delta x} + G(t)\overline{\delta u}$$

$$\dot{\overline{\delta x}} = \begin{bmatrix} (u_{1o} + C_T T_o + \frac{C_P u_{2o}}{\sqrt{T_o}} - \beta)/\ell & \lambda & (C_T Q_o - \frac{C_P u_{2o} Q_o}{2})/\ell \\ \beta/\ell & -\lambda & 0 \\ \frac{1}{M_c} & 0 & -a u_{2o} \end{bmatrix} \overline{\delta x} + \begin{bmatrix} \frac{Q_o}{\ell} & \frac{C_P Q_o}{\ell\ T_o} \\ 0 & 0 \\ 0 & -aT_o \end{bmatrix} \overline{\delta u} \tag{3.15}$$

The optimal feedback control is a 1 x 2 vector in δu_{1_o} and δu_{2_o}. From Eq. (1.11) the solution of these controls become:

$$\delta u_{1_o} = \frac{Q_o}{\ell} (P_{11}\delta Q + P_{12}\delta C + P_{13}\delta T) \tag{3.16}$$

$$\delta u_{2_o} = -\frac{C_P Q_o}{\ell\sqrt{T_o}} (P_{11}\delta Q + P_{12}\delta C + P_{13}\delta T) + aT_o(P_{31}\delta Q + P_{32}\delta C + P_{33}\delta T) \tag{3.17}$$

where the P_{ij} are typical elements of the matrix solution of the Ricatti matrix Eq. (1.13), P(t). Here, it is assumed that the control weighting matrix Q_2 is a 2 x 2 identity matrix. Again the error weighting matrix Q_1 is an arbitrary 3 x 3 diagonal matrix whose elements depend on the required accuracy.

The covariance matricies A(t) and R(t) must be determined by the statistical amplitudes of perturbations about all optimal trajectories.

These in turn are used in the solution of Eq. (1.22) to optimally predict the state variables.

The measurable state variables in this example are both temperature and power level. In all probability the power level in a nuclear rocket system will be measured logarithmically, but for digital synthesis linear measurement can be assumed.

The measurement matrix is thus:

$$M(t) = \begin{bmatrix} m_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & m_{33} \end{bmatrix} \tag{3.18}$$

where

$$m_{11} = m_{33} = 1$$

The four optimal feedback control schemes advanced in the previous chapter can easily be investigated for this problem digitally.

Chapter 4

DISCUSSION OF THE NUMERICAL METHODS

The first step in the numerical synthesis of the optimal closed loop problem was to determine values for the optimal state variables. In example 1, this was done by making a polynomial approximation as a function of time of a graphical representation of the optimal state variable, which was obtained from an analog computer solution of the state and auxiliary equations. Using eight boundary conditions along this trajectory, Crout's method of matrix reduction was applied to determine the coefficients for the polynomial expressions of n and $\dot{n}$. These polynomial expressions

$$n = 10.0 + At^2 + Bt^3 + Ct^4 + Dt^5 + Et^6 + Ft^7 + Gt^8$$

$$\dot{n} = 2At + 3Bt^2 + 4Ct^3 + 5Dt^4 + 6Et^5 + 7Ft^6 + 8Gt^7 \qquad (4.1)$$

were used to determine all optimal trajectories for 47 increments of time, with $\Delta t = 0.01$ seconds, and stored as reference data for the synthesis problem.

The Crout method[21], developed in 1941 for desk calculator by an electrical engineer, P. D. Crout, is particularly well-suited for solving simultaneous linear algebraic equations on the digital computer. Because both the recording of new arrays and the performing of repeated row operations at each intermediate stage of the reduction are not necessary, the Crout method is more efficient in terms of time and far less conducive

to gross error than the more widely used Gauss-Jordan method.

This method transforms the original matrix A into a triangular revised matrix A' by the following operations upon the elements of the matrix:

$$a'_{ij} = a_{ij} - \sum_{i=1}^{j-1} a'_{ik}\, a'_{kj} \quad (i \leq j)$$

$$a'_{ij} = \frac{1}{a'_{ii}} \left(a_{ij} - \sum_{k=1}^{j-1} a'_{ik}\, a'_{kj}\right) \quad (i > j) \tag{4.2}$$

where a_{ij} are typical elements of the A matrix.
The solution to the system is then calculated from the transformed matrix by back substitution from bottom to top according to the relationship

$$x_i = a'_{in} - \sum_{k=i+1}^{n} a'_{ik}\, x_k \tag{4.3}$$

These relationships are fully derived in reference 22, page 486.

The next step was to evaluate the typical elements P_{ij} and V_{ij} of Eqs. (1.13) and (1.22). These are systems of first order nonlinear differential equations. The usual method of attack on systems of differential equations has been the Milne Method or Runge-Kutta Method. The first is a predictor-corrector method involving the use of two quadrature formulas; the second is essentially an averaging method. Because of the size of the systems involved, a combination of two well-known methods, the Trapezoidal rule and the Newton-Raphson iterative method, is chosen instead to achieve the solutions.

The solution of the Ricatti-type matrix differential equation is a symmetric matrix. For example, six, rather than nine, simultaneous differential equations result from a 3 x 3 matrix solution.

These equations were then integrated numerically as a function of time using the Trapezoidal rule. The result is, instead of nonlinear differential equations, nonlinear algebraic equations. The familiar Newton-Raphson technique was chosen to solve these equations. This technique, in conjunction with the previously discussed Crout matrix reduction, is well adapted to the digital computer solution of this type of problem. The form of the integrated equations is, by the Trapezoidal rule,

$$P_{ij}(t_k) - P_{ij}(t_f) = -\Delta t \quad f^P_{ij} \frac{(t_f)}{2} + \sum_{\ell=1}^{k-1} f^P_{ij}(t_\ell) + \frac{f^P_{ij}(t_k)}{2} \tag{4.4}$$

$$V_{ij}(t_k) - V_{ij}(0) = \Delta t \quad f^V_{ij} \frac{(0)}{2} + \sum_{\ell=1}^{k-1} f^V_{ij}(t_\ell) + \frac{f^V_{ij}(t_k)}{2} \tag{4.5}$$

where f^P_{ij} and f^V_{ij} are the function forms of the P_{ij} and V_{ij} derivatives respectively. The minus sign is introduced in Eq. (4.4) since the P_{ij} are integrated backwards in time. The result can be expressed as the system of equations of the form

$$f_n(q_{ij}) = 0 \tag{4.6}$$

where the q_{ij} refer to either P_{ij} or q_{ij}

$$f_n(P_{ij}) = P_{ij}(t_k) + I^P_\ell \quad k + \Delta t \; f^P_{ij}(t_k)/2 \tag{4.7}$$

$$f_n(V_{ij}) = V_{ij}(t_k) - I^V_\ell \quad k - \Delta t \; f^V_{ij}(t_k)/2 \tag{4.8}$$

where

$$I^P_\ell \quad k = \Delta t \quad \frac{f^P_{ij}(t_f)}{2} + \sum_{\ell=1}^{k-1} f^P_{ij}(t_\ell) \tag{4.9}$$

and

$$I^V_{\ell}\ k = \Delta t\ \ f^V_{ij}\frac{(0)}{2} + \sum_{\ell=1}^{k-1} f^V_{ij}(t_{\ell}) \tag{4.10}$$

The Newton-Raphson Method is one of successive iterations. For the digital computer solution the best method of obtaining first approximations for the kth time period was the Runge-Kutta technique as a predictor based on the final result of the k-1th time period.

$$P_{ij}(t_k)_o = P_{ij}(t_{k-1}) + \frac{1}{6}(R^P_{1_{ij}} + 2R^P_{2_{ij}} + 2R^P_{3_{ij}} + R^P_{4_{ij}}) \tag{4.11}$$

where

$$R^P_{1_{ij}} = -\Delta t\ f^P_{ij}(P_{ij})$$

$$R^P_{2_{ij}} = -\Delta t\ f^P_{ij}(P_{ij} + R_{1_{ij}}/2)$$

$$R^P_{3_{ij}} = -\Delta t\ f^P_{ij}(P_{ij} + R_{2_{ij}}/2)$$

$$R^P_{4_{ij}} = -\Delta t\ f^P_{ij}(P_{ij} + R_{3_{ij}})$$

and

$$V_{ij}(t_k)_o = V_{ij}(t_{k-1}) + \frac{1}{6}(R^V_{1_{ij}} + 2R^V_{2_{ij}} + 2R^V_{3_{ij}} + R^V_{4_{ij}})$$

where

$$R^V_{1_{ij}} = \Delta t\ f^V_{ij}(V_{ij}) \tag{4.12}$$

$$R^V_{2_{ij}} = \Delta t\ f^V_{ij}(V_{ij} + R^V_{1_{ij}}/2);$$

$$R^V_{3_{ij}} = \Delta t\ f^V_{ij}(V_{ij} + R^V_{2_{ij}}/2)$$

$$R^V_{4_{ij}} = \Delta t\ f^V_{ij}(V_{ij} + R^V_{3_{ij}})$$

Again the minus sign appears in $R_{n_{ij}}$ because of backward integration. The Newton-Raphson technique is a linearizing iterative technique for determining small deviations in the q_{ij} from their true values as determined from the following equations

$$\frac{\partial f_1}{\partial q_{11}} \Delta q_{11} + \frac{\partial f_1}{\partial q_{12}} \Delta q_{12} + \cdots + \frac{\partial f_1}{\partial q_{mm}} \Delta q_{mm} + f_1(t_1) = 0$$

$$\frac{\partial f_n}{\partial q_{11}} \Delta q_{11} + \frac{\partial f_n}{\partial q_{12}} \Delta q_{12} + \cdots + \frac{\partial f_n}{\partial q_{mm}} \Delta q_{mm} + f_n(t_{k-1}) = 0 \qquad (4.13)$$

where

$$n = m(m + 1)/2$$

The increments Δq_{ij} are

$$\Delta q_{ij} = q_{ij_\ell} - q_{ij_{\ell-1}}$$

where ℓ refers to the iteration number.

Eqs. (4.13) are linear simultaneous algebraic equation in q_{ij} which can be solved by Crout's method. These are then the new values for the next iteration. Convergence by this method is generally rapid. Care must be taken that all time varying entities are arranged in the proper order for each set of calculations.

The digital computer program is given in the appendix and is entitled "Solution of the Time-Varying Matrix Ricatti Differential Equation"[23] and will be published at a later date. Figures 9 and 10 show flow charts for the computer code.

All random disturbances were generated digitally by means of a "canned" number generator function within the IBM 7072 systems tape.

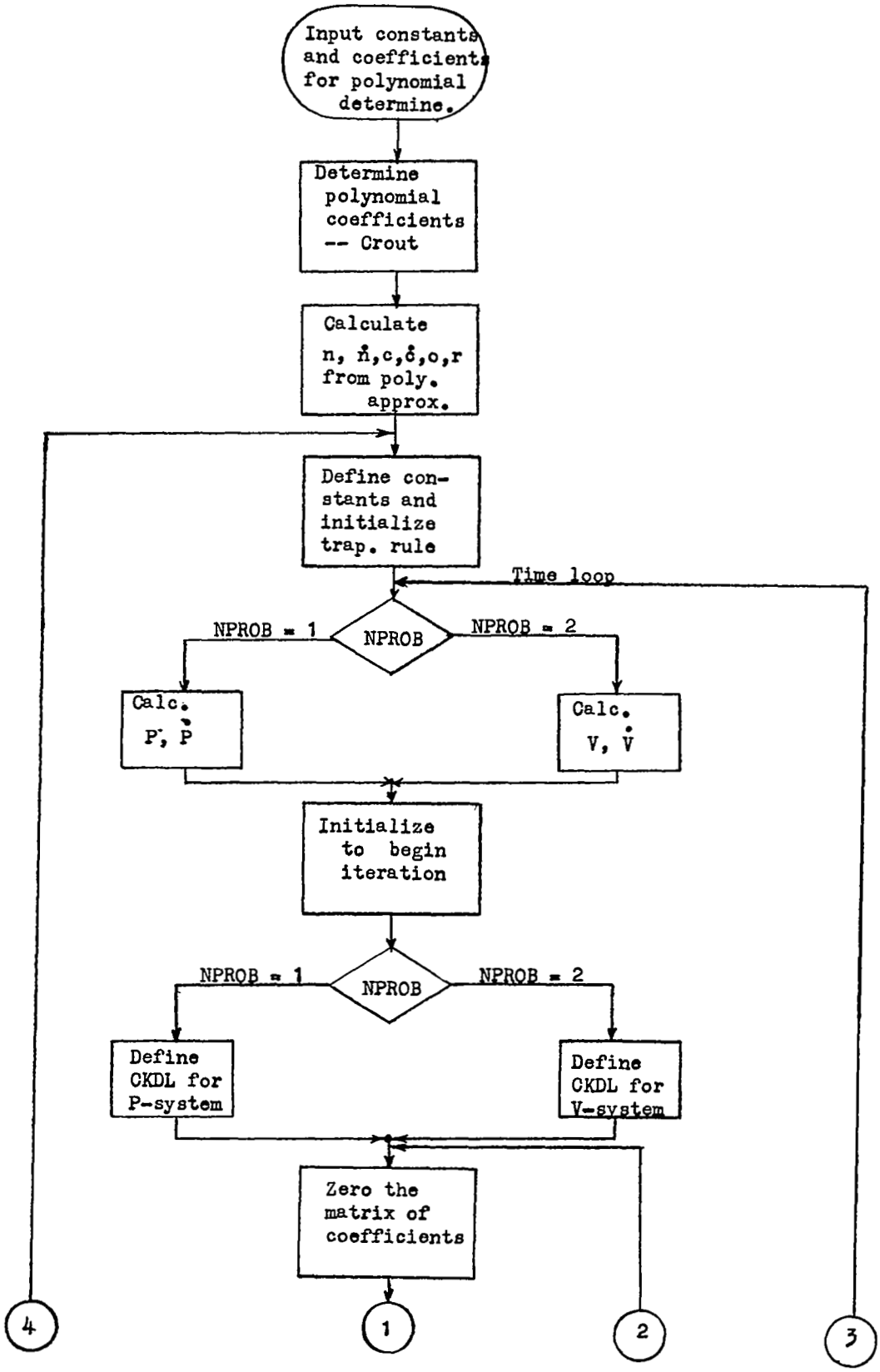

Input constants and coefficients for polynomial determine.
Determine polynomial coefficients -- Crout
Calculate n, ṅ, c, ċ, o, r from poly. approx.
Define constants and initialize trap. rule
Time loop
NPROB
NPROB = 1
NPROB = 2
Calc. Ṗ, P
Calc. V̇, V
Initialize to begin iteration
NPROB
NPROB = 1
NPROB = 2
Define CKDL for P-system
Define CKDL for V-system
Zero the matrix of coefficients
4
1
2
3

Figure 9

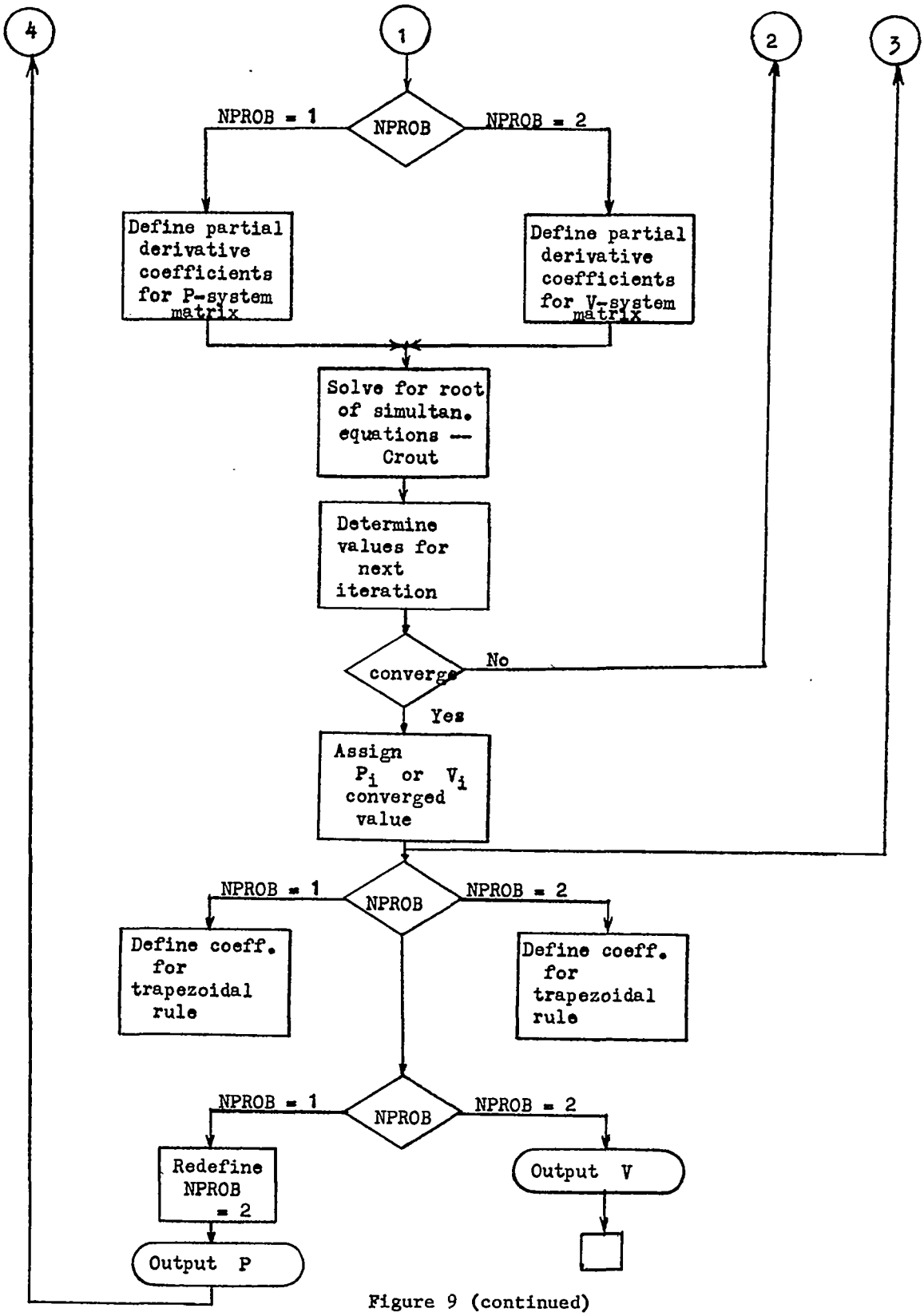


Figure 9 (continued)

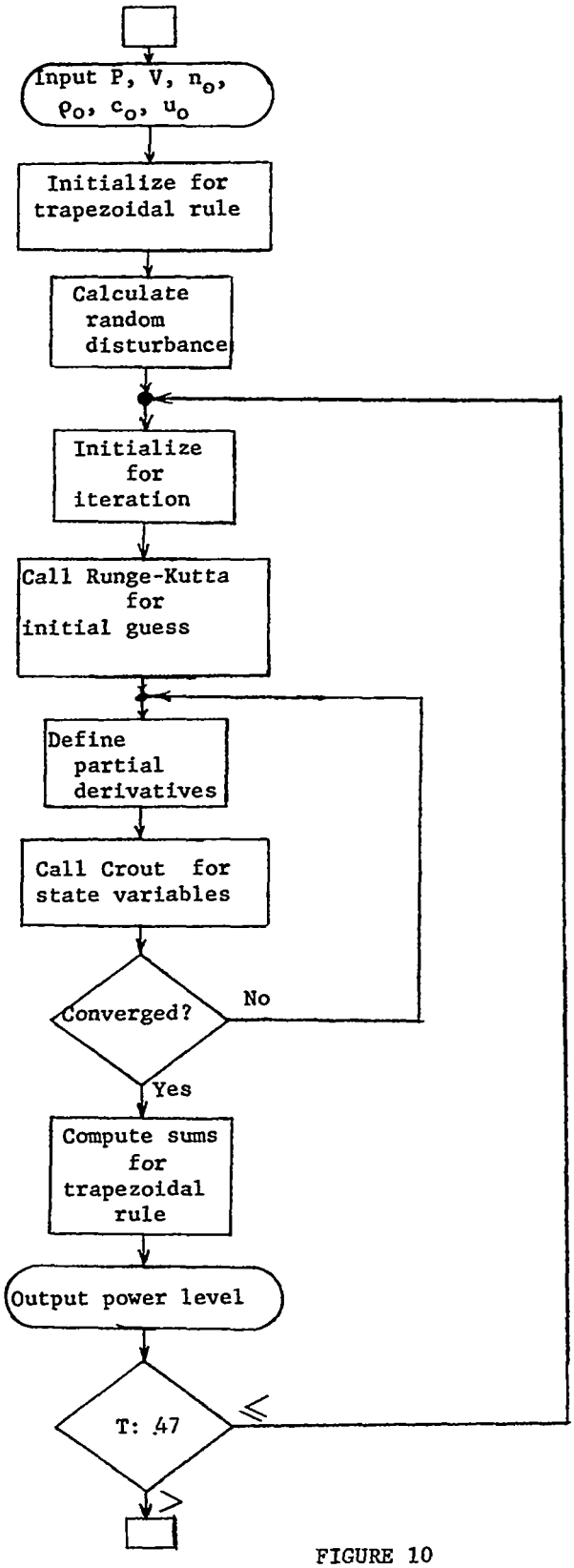

Input P, V, n_o, ρ_o, c_o, u_o
Initialize for trapezoidal rule
Calculate random disturbance
Initialize for iteration
Call Runge-Kutta for initial guess
Define partial derivatives
Call Crout for state variables
Converged?
No
Yes
Compute sums for trapezoidal rule
Output power level
T: 47

FIGURE 10

The random disturbances were assumed to have zero mean with a Gaussian distribution given by the well known bell-shaped formula,

$$\phi = \exp\ (-RD^2/2\sigma_{RD}{}^2)/(\sqrt{2\pi}\ \sigma_{RD}) \tag{4.15}$$

where ϕ is the ordinate of the normal curve, RD, is the random disturbance. For digital computation all random numbers from 0.0 to 1.0 were set equal to the exponential of Eq. (4.15).

$$x = \exp\ (-RD^2/2\sigma_{RD}{}^2) \tag{4.16}$$

The random disturbance in terms of the random number becomes

$$RD = \ (2\ \ln(\frac{1}{x}))^{1/2}\ \sigma_{RD} \tag{4.17}$$

Since the random disturbance must have zero mean negative values of RD must be equally as probable as positive values. On a random number generator the random numbers have a rectangular distribution; hence numbers greater than 0.5 are as likely as numbers less than 0.5. For x greater than 0.5 it was assumed that RD was positive; for x less 0.5, RD was negative. Eq. (4.17) was then modified

$$RD = \begin{cases} +\ (2\ \ln(\frac{1}{2|x-.5|}))^{1/2}\ \sigma_{RD} & 0.5 < x \leq 1.0 \\ -\ (2\ \ln(\frac{1}{2|x-.5|}))^{1/2}\ \sigma_{RD} & 0 \leq x \leq 0.5 \end{cases} \tag{4.18}$$

All variances were determined to constrain the amplitudes of the perturbations.

For synthesis by digital computation straightforward solution of the equations outlined in Chapter 1 was performed for discrete time

intervals. Where the solution of sets of simultaneous nonlinear differential equations was required, each was integrated by the Trapezoidal rule and then Runge-Kutta approximate predictions and Newton-Raphson iterations made to obtain the final accuracy.

Chapter 5

AREAS FOR CONTINUING STUDY

A major problem in the area of optimal control arises in the solution of the optimal state transition trajectories. Techniques in use and under study for determination of such trajectories have been classified as direct and indirect. The direct method is characterized by a systematic scheme to search for the optimum. The method of steepest decent, or gradient method, is an example. Indirect methods include Pontryagin's Maximum Principle,and calculus of variations which result in a set of differential equations, the boundary conditions being incomplete on both ends of the solution interval. These are classified as two point boundary value problems.

Knapp[25] has introduced a technique which employs the gradient method for solving the two point boundary value problem thus combining both direct and indirect methods. The approach has been successfully used to solve six simultaneous nonlinear differential equations by the computer. An effort is being made to develop a code which will solve the two point boundary value problems of interest in nuclear reactor dynamics. The result would make possible the synthesis of a large number of unsolved optimal control problems in this field.

Presently, only the one problem has been investigated using the Linear Optimal Stochastic Control Theory. Very little has been done in the way of parameter variations. However, with the existence of the

computer techniques investigation of many parameter changes is now possible for the problem of example 1 with little additional effort.

With the completion of the first example an extensive study will be made of optimal closed loop control for nuclear rocket engine start-up (shut-down). A clear definition of the optimal control law will be determined. This study will include parameter variations. A close liason will be maintained with the Los Alamos Scientific Laboratory to keep up to date with current problems and to obtain valuable advice.

Linear Optimal Stochastic Control Theory has the disadvantage that the precise optimal (or nominal) trajectory must be previously known to determine the optimal feedback process. Thus different feedback gains must be determined for each transition trajectory anticipated. For nuclear-rocket engine start-up this does not present a problem, but for most applications this becomes a task. Optimal closed loop control by other synthesis techniques will be compared with this one.

REFERENCES

1. R. H. Mohler, "Optimal Control of Nuclear Reactor Systems," ANS Annual Meeting TRANSACTIONS, pp. 58-59, June 1964.

2. Z. R. Rosztoczy, A. P. Sage and L. E. Weaver, "Application of Pontryagin's Maximum Principle to Flux-State Changes in Nuclear Reactors," TID 7662 AEC Symposium Series 2, Reactor Kinetics and Control, 1964.

3. R. J. Hermsen and T. J. Higgins, "Optimal Design of Nuclear Reactor Control Systems by Use of Pontryagin's Maximum Principle, Conference paper presented at IEEE Salt Lake City meeting, April, 1964.

4. C. N. Shen and F. G. Haag, "Application of Optimal Control to Nuclear Reactor Start-Up," IEEE - TYGNS, Vol. NS-11, No. 2, pp. 1-8, April, 1964.

5. L. S. Pontryagin, et.al., The Mathematical Theory of Optimal Processes, Interscience Publishers, John Wiley and Sons, New York, 1964.

6. N. Wiener, The Extropolation, Interpolation and Smoothing of Stationary Time Series, John Wiley and Sons, New York, 1949.

7. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, Journal of Basic Engineering, pp. 35-45, March, 1960.

8. R. E. Kalman, and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Trans. ASME, Journal of Basic Engineering, pp. 95-108, March, 1961.

9. J. V. Breakwell, J. L. Speyer, and A. E. Bryson, "Optimization and Control of Nonlinear Systems Using the Second Variation," J.S.I.A.M. Control, Series A, Vol. 1, No. 2, 1963.

10. J. S. Tyler Jr., "The Characteristics of Model-Following Systems as Synthesized by Optimal Control," Proceedings of IEEE, J.A.C.C. Session I paper 7, San Francisco, June, 1964.

11. F. Tung, "Linear Control Theory Applied to Interplanetary Guidance," IEEE Trans. on Automatic Control, Vol. AC-9, pp. 82-90, January, 1964.

12. G. L. Smith, "Multivariable Linear Filter Theory Applied to Space Vehicle Guidance," presented at SIAM meeting on Control, Cambridge, Massachusetts, October, 1962.

13. T. L. Gunkel, III and G. F. Franklin, "A General Solution for Linear, Sampled Data Control," Trans. ASME, Journal of Basic Engineering, Vol. 85-D, pp. 197-201, June, 1963.

14. H. Cox, "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," IEEE Trans. on Automatic Control, Vol. AC-9, No. 1, January, 1964.

15. C. S. Weaver, "Estimating the Output of a Linear Discrete System With Gaussian Inputs," IEEE Trans. on Automatic Control, Vol. AC-8, No. 4, October, 1963.

16. J. J. O'Donnell, "Finite-Time Filtering: One Dimensional Case," IEEE Trans. on Automatic Control, Vol. AC-9, No. 1, January, 1964.

17. R. Cosaert, "On the Estimation of the Plant State, Plant Parameters, and External Disturbances From Observations of the Output of a Control System," Proceedings of IEEE, JACC San Francisco, June, 1964.

18. H. P. Smith and A. H. Stenning, "Dynamics and Control of Nuclear Rocket Engines," Nuclear Science and Engineering, Vol. 11 No. 1 (1961.)

19. R. R. Mohler and J. E. Perry, Jr., "Nuclear Rocket Engine Control," Nucleonics, Vol. 19, No. 4, April, 1961.

20. Bussard and DeLauer, Nuclear Rocket Propulsion, McGraw Hill Book Co., Inc., New York, 1958.

21. R. R. Mohler, "Optimal Control of Nuclear-Reactor Processes," Ph.D. Thesis (Michigan 1964) unpublished.

22. F. B. Hildebrand, Introduction to Numerical Analysis, McGraw Hill Book Co., Inc., pp. 82, 429, 434, 447, 451, 453, New York, 1956.

23. P. A. Secker, Jr. and S. T. Murray, Computer Code, "Solution of the Time Varying Matrix Ricatti Differential Equation," University of Arizona, August 1964, unpublished.

24. K. J. Nielson, Methods in Numerical Analysis, Macmillan, New York, 1956.

25. C. H. Knapp and P. A. Frost, "Determination of Optimal Control and Trajectories Using the Maximum Principle in Association with a Gradient Technique," Proceedings of IEEE, Joint Automatic Control Conference, Stanford University, June, 1964.

SECTION IV

LIMITS OF VALIDITY FOR SOME APPROXIMATIONS IN REACTOR DYNAMICS

One of the problems of nuclear rocket dynamics which may be treated by optimization theory is the problem of minimum-time start-up. This problem requires mathematical models for the neutron dynamics, for the effects of temperature on reactivity, and for the necessary constraints such as limitations on maximum thermal stresses. An important simplification results if the response of the reactor to a constant rate of reactivity increase can be described by a simple approximation.

The reactor dyanmics equations have simple solutions only when the reactivity is not an explicit function of time. Among the well-known approximate solutions are the "prompt-jump" approximation[1,2] (hereinafter called PJ), in which the prompt neutron lifetime ℓ is neglected, and the "rapid-rate" approximation[3,4] (hereinafter called RR), in which the delayed-neutron decay constant λ is neglected. The combination of these two approximations yields, in the usual notation,

$$n/n_o = \beta/(\beta - \rho) \tag{1}$$

where n_o is the steady-state neutron density for $t<0$ ($\rho = 0$); this result may also be derived from simple physical considerations[5]. In the special case of a "ramp input" of reactivity ($\rho = \gamma t$), Eq. (1) becomes[4]

$$n/n_o = \beta/(\beta - \gamma t) \tag{2}$$

The purpose here is to determine the conditions under which Eqs. (1) and (2) are useful approximations.

Assuming one group of delayed neutrons, the dynamic equations are

$$\dot{n} = \frac{\rho - \beta}{\ell} n + \lambda c, \tag{3}$$

$$\dot{c} = \frac{\beta}{\ell} n - \lambda c. \tag{4}$$

Eliminating c yields

$$\ell\ddot{n} + (\beta - \rho + \lambda\ell)\dot{n} - (\dot{\rho} + \lambda\rho) n = 0. \tag{5}$$

If we neglect $\dot{n}$ in Eq. (3), that is, if

$$\ell\dot{n}/n \ll |\rho - \beta| \tag{6}$$

we obtain

$$(\beta - \rho)\dot{n} - (\dot{\rho} + \lambda\rho)n = 0. \tag{7}$$

This is also obtained by setting $\ell = 0$ in Eq. (5). Eq. (7) may be solved when $\rho = \gamma t$, yielding

$$n/n_o = \left[\beta/(\beta - \gamma t)\right]^{1 + \lambda\beta/\gamma} e^{-\lambda t} \tag{8}$$

which is the well-known ramp response in the PJ approximation. Eq. (2) results if $\lambda = 0$ in Eq. (8); more basically, Eq. (1) may be derived directly from Eq. (3) by neglecting $\dot{n}$ and replacing λc by its steady-state value $\beta n_o/\ell$.

To determine the range of validity for the PJ approximation, compare Eqs. (6) and (7) to obtain $(\beta - \rho)^2 >> \ell(\dot{\rho} + \lambda\rho)$. Since the

main concern is the closeness of approach to prompt critical, the condition for validity of the PJ approximation may be expressed as

$$\beta - \rho \gg \sqrt{\ell(\dot{\rho} + \lambda\beta)} \tag{9}$$

For fast ramps, this becomes

$$\beta - \rho \gg \sqrt{\ell\,\dot{\rho}} \tag{10}$$

while for slow ramps, Eq. (9) reduces to

$$\beta - \rho \gg \sqrt{\ell\lambda\beta} \tag{11}$$

Eq. (11), written as

$$\beta - \rho > 2\sqrt{\ell\lambda\beta}$$

is cited by Cohen[6] as the criterion for validity of the prompt-jump approximation. Its usefulness is restricted to slow ramps.

For the RR approximation, set $\lambda c = \beta n_o/\ell$ in Eq. (3), differentiate with respect to time, and compare the result with Eq. (5). The significant requirement is then seen to be $\dot{\rho} \gg \lambda\rho$. For a ramp, this yields

$\rho \ll \gamma/\lambda$, or

$$t \ll 1/\lambda \tag{12}$$

as expected. A similar result is obtained by expanding Eqs. (2) and (8) in powers of t and comparing coefficients of t^2. The validity of Eq. (1) is then governed by Eqs. (9) and (12) together.

An example is shown in Fig. 1, for which $\gamma = 0.1$ dollar/sec, $\lambda = 0.1\ \text{sec}^{-1}$, and $\beta/\ell = 100\ \text{sec}^{-1}$. The PJ curve is from Eq. (8). The exact solution of Eq. (5) for $\rho = \gamma t$ may be expressed in terms of hypergeometric functions, but if $\lambda\beta/\gamma$ is an integer, a simpler form

involving the error function results[7,8]. (See also References 9, 10, and 11.) Here $\lambda\beta/\gamma = 1$, and the exact solution for $n(0) = n_o$ and $\dot{n}(0) = 0$ is

$$n/n_o = \frac{\beta^2}{\gamma\ell}\left\{ e^{-\lambda t} - \frac{\beta-\lambda\ell-\gamma t}{\beta} \exp\left(\frac{\gamma}{2\ell} t^2 - \frac{\beta}{\ell} t\right)\left[1 + \beta\sqrt{\pi/2\gamma\ell}\ \exp\frac{(\beta - \lambda\ell)^2}{2\gamma\ell}\left(\operatorname{erf}\frac{\beta - \lambda\ell}{\sqrt{2\gamma\ell}} - \operatorname{erf}\frac{\beta - \lambda\ell-\gamma t}{\sqrt{2\gamma\ell}}\right)\right]\right\} \quad (13)$$

For t not near β/γ (provided $t < \beta/\gamma$) the error functions have large arguments, and the asymptotic forms yield

$$n/n_o = (\beta^2/\gamma\ell)e^{-\lambda t}\left(\frac{1}{x^2} - \frac{1.3}{x^4} + \frac{1.3.5}{x^6} - \ldots\right)$$

where $x = (\beta - \lambda\ell - \gamma t)/\sqrt{\gamma\ell}$. This becomes the PJ approximation for $\lambda\beta/\gamma = 1$ if x is large and if $\lambda\ell$ is neglected compared to β.

The RR curve is computed from[3,7]

$$n/n_o = \left[1 + \beta\sqrt{\pi/2\gamma\ell}\ \exp\frac{\beta^2}{2\gamma\ell}\left(\operatorname{erf}\frac{\beta}{\sqrt{2\gamma\ell}} - \operatorname{erf}\frac{\beta - \gamma t}{\sqrt{2\gamma\ell}}\right)\right] \cdot \exp\left(\frac{\gamma}{2\ell} t^2 - \frac{\beta}{\ell} t\right) \quad (14)$$

For t not near β/γ, this becomes

$$n/n_o = \frac{\beta}{\beta - \gamma t}\left(1 - \frac{1}{y^2} + \frac{1.3}{y^4} - \frac{1.3.5}{y^6} + \ldots\right)$$

where $y = (\beta - \gamma t)/\sqrt{\gamma\ell}$; for large y this becomes Eq. (2). Note that Eq. (14) does not follow from Eq. (13) as $\lambda \to 0$ since Eq. (13) requires $\lambda\beta/\gamma = 1$.

The foregoing considerations have not included mention of one important characteristic of the PJ approximation, namely its failure for very small values of t. For a ramp input starting from a steady state at $t = 0$, Eq. (3) requires that $\dot{n}(0) = 0$. The PJ approximation yields a discontinuity in $\dot{n}$ at $t = 0$; however, this represents a transient which vanishes in a time comparable with ℓ/β and which has an effect so small that it is not observable in Fig. 1.

To illustrate the criteria, replace Eq. (9) by

$$\beta - \rho > \sqrt[3]{\ell(\dot{\rho} + \lambda\beta)}. \tag{15}$$

For the numerical example, this is $1 - \rho/\beta > 0.135$, or $t < 8.7$ sec. If Eq. (12) is replaced by

$$t < 1/3\lambda, \tag{16}$$

we have $t < 3.3$ sec. As verified by Fig. 1, Eq. (12) is dominant in determining the validity of Eq. (2) in this example; for large ramp rates, Eq. (9) will dominate.

The criterion for validity of the PJ approximation given by Eq. (15) may be displayed graphically for ramp inputs by plotting contours of constant ρ_m as in Fig. 2. The co-ordinate axes are β/ℓ and $\dot{\rho}/\beta = \gamma/\beta$, and ρ_m is given by

$$\beta - \rho_m = \sqrt[3]{\ell(\dot{\rho} + \lambda\beta)}. \tag{17}$$

Hence ρ_m is the maximum reactivity for a given ℓ and γ for which the PJ approximation is valid within the limit set by Eq. (15); i.e., the PJ approximation is valid in a region to the right of a given contour for reactivities at least as large as that on the contour defining the region.

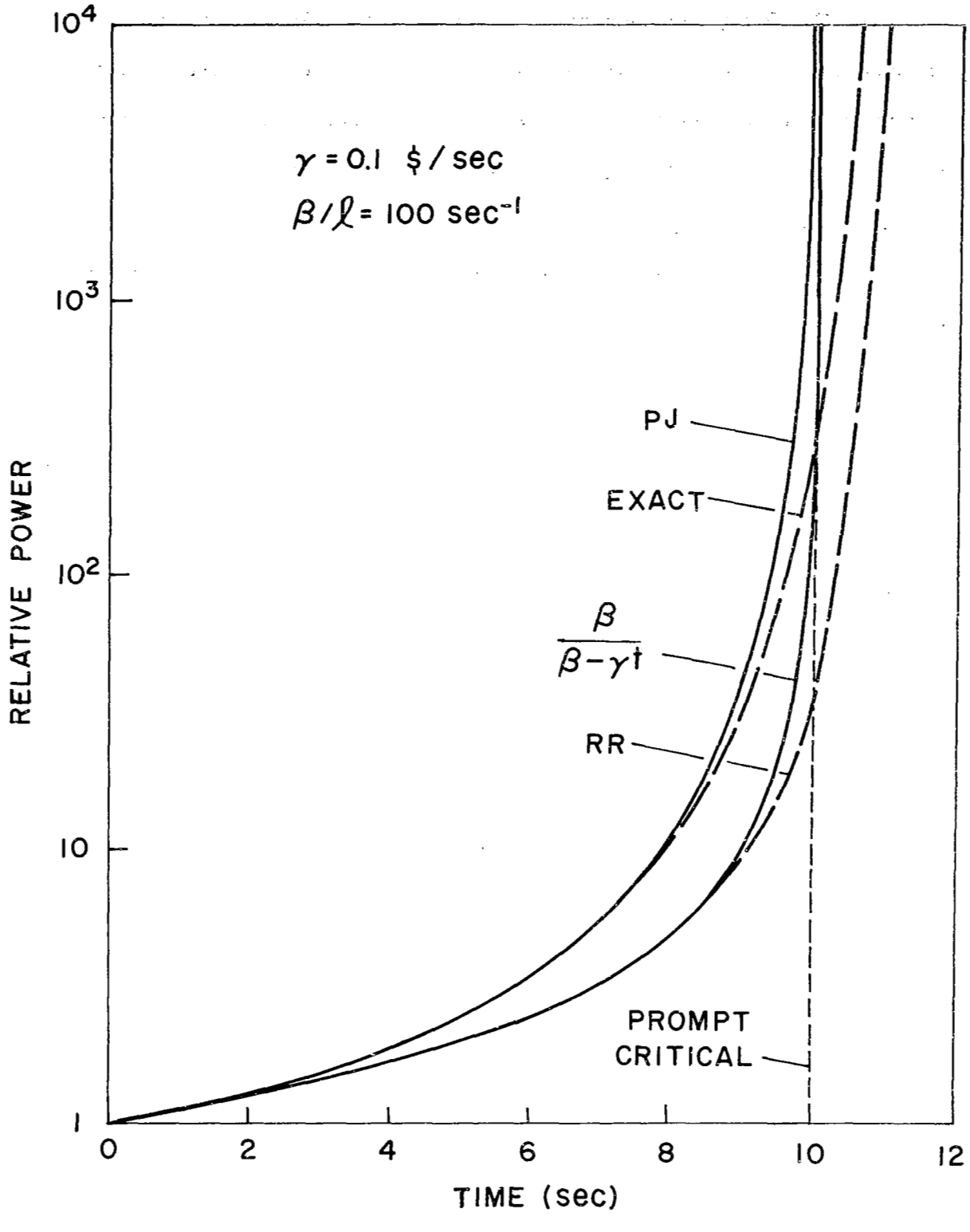


Fig. 1. Comparison of Approximations: Response of a Reactor to a Ramp Input of Reactivity.

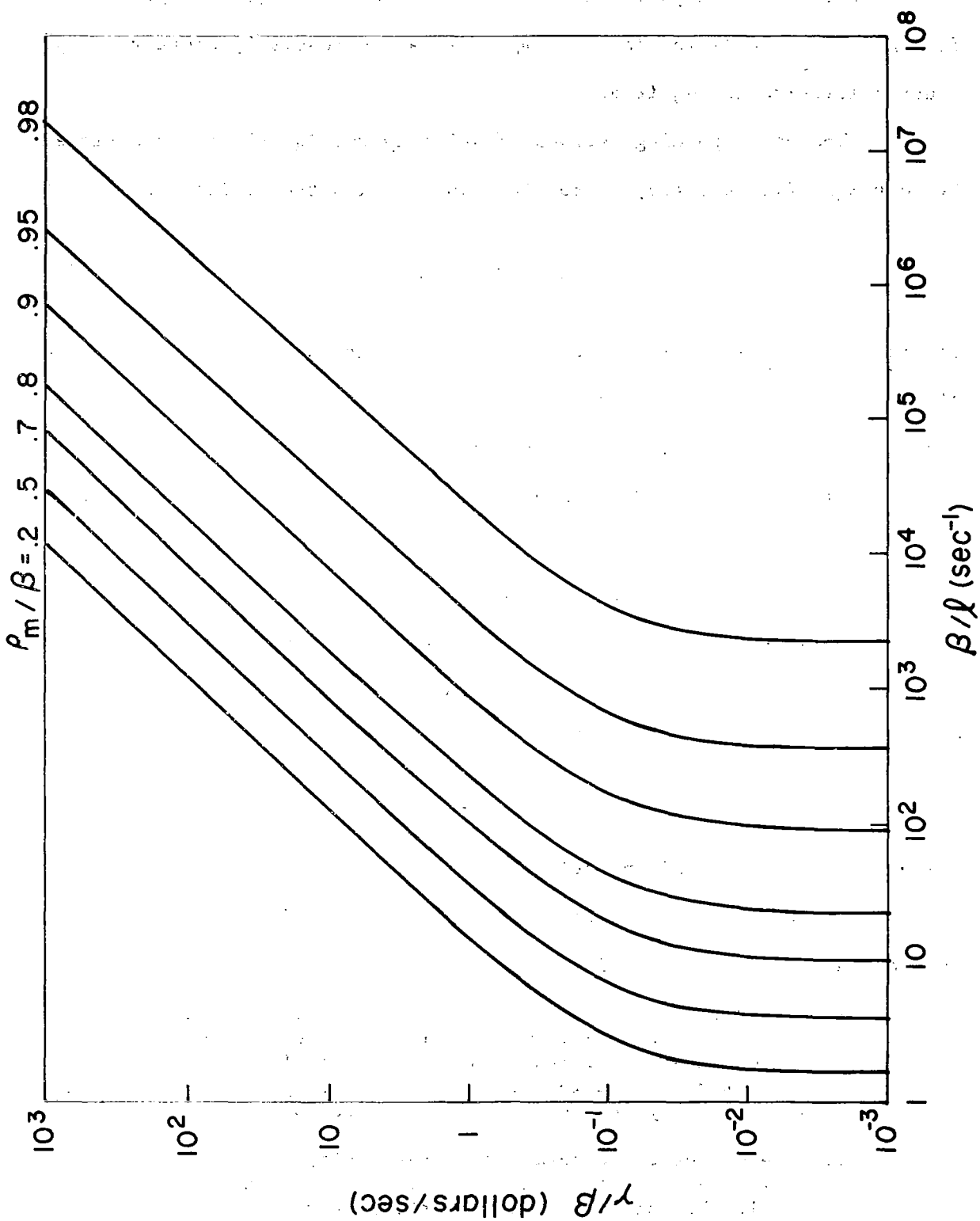


Fig. 2. Regions of Validity for the Prompt-Jump Approximation.

Note that for slow ramps the limits are independent of γ, as expected from Eq. (11), while for larger ramp rates a progressively shorter prompt neutron lifetime is required.

The corresponding regions of validity for Eq. (2) are restricted by both Eqs. (15) and (16). From Eq. (16) we find the limiting reactivity

$$\rho_m = \gamma/3\lambda. \tag{18}$$

The maximum reactivity for which Eq. (2) is valid is therefore the smaller of the two values given by Eqs. (17) and (18). The modified regions are shown in Fig. 3, reflecting the fact that Eq. (2) is not useful if the ramp rate is too small. This is an obvious consequence of the assumption of a constant production rate for delayed neutrons.

The results are easily extended to start-up calculations in which the initial steady state is maintained by an extraneous source of neutrons. In this case, Eq. (8) is replaced by a much more complicated form[1], but Eq. (1) has a simple extension[4]:

$$n/n_o = (\beta - \rho_o)/(\beta - \rho),$$

where ρ_o is the initial (negative) reactivity.

Further study is necessary before the preceding is incorporated into an investigation of optimization of fast start-ups; in particular:

1) The approximate solution, Eq. (2) or Eq. (8), must be terminated before prompt critical because of the obvious divergence, and a means must be devised for carrying an approximate solution smoothly past this point and matching it to an asymptotic form of the exact solution.

2) The effect of temperature on reactivity must be incorporated; this need not be included until the later stages of a start-up if the initial level is sufficiently low.

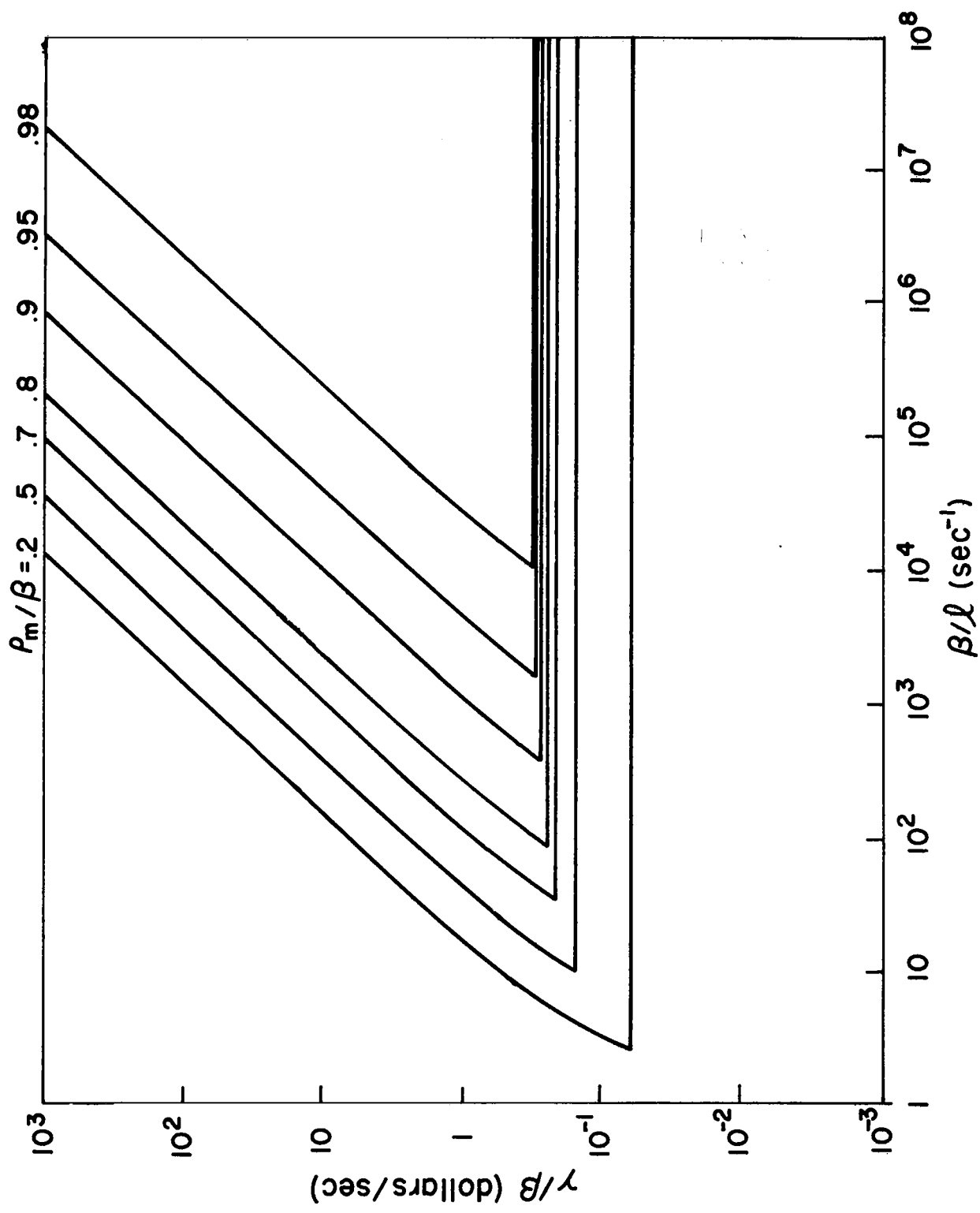


Fig. 3. Regions of Validity for Eq. (2).

The problem of matching approximate solutions across prompt critical has been investigated by MacPhee[12] in studies of reactor accidents. Numerous rough calculations have been made which provide conservative over-estimates for accident studies. It is hoped that further study will yield approximate solutions which are more suitable for the fast start-up optimization problem.

REFERENCES

1. Syrett, J. J., "The Variation of Power in a Fast Reactor for a Linear Increase of Reactivity," AERE-RP/M-47 (August 16, 1954).

2. Bethe, H. A., "Reactor Safety and Oscillator Tests," APDA-117 (October 15, 1956).

3. Schultz, M. A., "Control of Nuclear Reactors and Power Plants," McGraw-Hill Book Co., Inc., 1st ed., p. 42 (1955).

4. Soodak, H. (ed.), "Reactor Handbook, Physics," Interscience Publishers, Vol. III, Part A, pp. 250-6 (1962).

5. Kovanic, P., "An Approximate Solution of Reactor Kinetics Equations," Proc. Geneva Conf. 11 293 (1958).

6. Cohen, E. R., "Control Problems in Nuclear Reactors," in "Computer Control Systems Technology," McGraw-Hill Book Co., Inc., p. 507 (1961).

7. Smets, H. B., "Reactor Dynamics at Low Power," Proc. Geneva Conf. 11 237 (1958).

8. Smets, H. B., "Exact Solution of the Reactor Kinetics Equations for Linear, Exponential, or Reciprocal Reactivity Dependence on Time," Bull. Acad. Roy. Belg. Cl. Sci. 45 256 (1959).

9. Coveyou, R. R., and Mulliken, T. W., "Solution of the Equation $\ddot{N} + (a + bt)\dot{N} + (c + dt)N = 0$," AECD-2407 (December 9, 1948).

10. Wallach, S., "Solutions of the Pile Kinetic Equations When the Reactivity is a Linear Function of the Time," WAPD-13 (November 30, 1950).

11. Hurwitz, H., Jr., "Approximate Analysis of Reactor Start-Up Incidents," Nucl. Sci. and Eng. 6 11 (1959).

12. MacPhee, J., "Approximate Solutions of the Reactor Kinetic Equations for Ramp Inputs," Nucl. Sci. and Eng. 7 33 (1960).